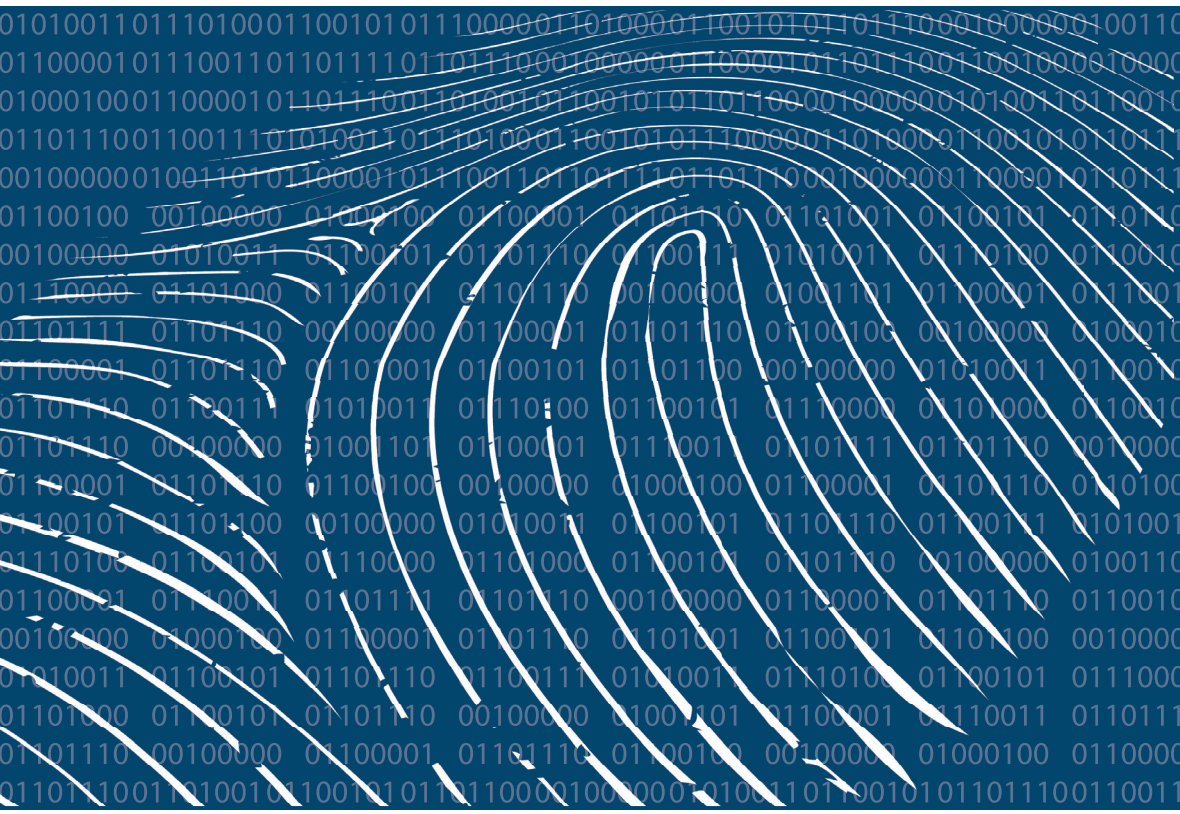
IALS INSTITUTE OF ADVANCED LEGAL STUDIES
SCHOOL OF ADVANCED STUDY UNIVERSITY OF LONDON



Electronic Evidence

Fourth edition

Editors: Stephen Mason and Daniel Seng

Electronic Evidence

OBserving Law – IALS Open Book Service for Law

First published by LexisNexis Butterworths, 2007

Second edition published by LexisNexis Butterworths, 2010

Third edition published by LexisNexis Butterworths, 2012

Fourth edition published by the Institute of Advanced Legal Studies for the SAS Humanities Digital Library, School of Advanced Study, University of London, 2017

ISBN 978-1-911507-07-9 (PDF edition)

This book is also available online at http://ials.sas.ac.uk/digital/humanities-digital-library/observing-law-ials-open-book-service-law.

Institute of Advanced Legal Studies

University of London

Charles Clore House

17 Russell Square

London WC1B 5DR

http://ials.sas.ac.uk

Electronic Evidence

Fourth edition

Stephen Mason and Daniel Seng

of the Middle Temple, Barrister

Associate Professor, National University of Singapore

Editors

Contributors

Dr George R. S. Weir
Computer and Information Sciences, University of Strathclyde

Burkhard Schafer
Professor of Computational Legal Theory, School of Law, University of Edinburgh

Chris Gallavin
Professor of Law and Deputy pro-Chancellor, Massey University

Dr Allison Stanfield
Managing Director and Solicitor, SG Legal Services Pty Ltd

Alisdair Gillespie
Professor of Criminal Law and Justice and Head of the Law School, Lancaster University

Andrew Sheldon, MSc
Evidence Talks Limited

Hein Dries
Vigilo Consult

A note on our Creative Commons licence, free legal resources on the Internet and citations of websites

For the first time, the authors, the editors and the publisher have collectively decided to make available the fourth edition of Electronic Evidence under a Creative Commons licence. We made this carefully considered decision because we want to promote a better understanding of electronic evidence, and wish to facilitate the greater accessibility and availability of our combined scholarship. We commend the Institute of Advanced Legal Studies, University of London, for its strong and continued support for academic education, learning and scholarship and the advancement of knowledge.

Most readers familiar with the common law will be aware of some of the free legal sources on the Internet. For the uninitiated, the World Legal Information Institute (www.worldlii.org) is a good start. Many of the more recent cases cited in this book, but by no means all, are available on the various independent jurisdiction-specific web sites that are linked to the World Legal Information Institute, which in turn is coordinated by the Australasian Legal Information Institute (www.austlii.edu.au), the first of its kind. Note also The Free Access to Law Movement (www.falm.info). Additional links can be found on any university library web site, including the website of the Institute of Advanced Legal Studies, London. It must be emphasized that the free sources of case law that are available are not comprehensive.

Readers will be familiar with the changing nature of URLs. Every effort has been made to ensure, where a URL is given, that it was live at the time of publication.

References have been made to Wikipedia on the basis that this source is relatively accurate for information of a technical nature. Readers will be aware that these pages are open to being up-dated and changed regularly. Although it is sometimes customary to provide the date a page was last viewed on the Internet, it is taken as a given that the reader does not need this information, given the dynamic nature of the Internet.

Errors and Omissions

While we, our authors and the publisher have tried hard to ensure all typographical and other errors have been corrected, we are aware that we might have missed some. For this reason, we will be delighted if you let us know if you notice an error. In addition, if you detect any relevant case law, legislation, guidelines or reports that we have missed, we will appreciate it if you inform us of any helpful and pertinent materials.

This edition is dedicated to the memory of a generous and distinguished man

Jon Bing

(30 April 1944 – 14 January 2014)

Professor dr juris – author

Senter for rettsinformatikk, Institutt for privatrett

Det juridiske fakultet

Universitetet i Oslo

Also to the memory of a man with a warm heart

Harald Hjort

(19 September 1955 – 12 July 2016)

One of the authors of the chapter on Norway in

International Electronic Evidence (2008)

Contents

Business records

Judge Nuri Efendi looked over his spectacles. 'Now we have covered the main matters to be dealt with in this case management conference, you may address the business records point, Mr Ayarcı.'

Mr Halit Ayarcı stood up. 'Your honour, thank you. My learned friend intends to submit a number of spreadsheets into evidence. There are problems with this. The first of which is that he only intends to submit print-outs of the spreadsheet application or program, whatever our technical friends consider a spreadsheet to be. My learned friend has declined to provide copies to the defence in electronic form. My application is for the prosecution to provide copies of the relevant spreadsheets in electronic form.'

Mr Hayri İrdal stood up. Mr Halit Ayarcı sat down.

'Your honour, I must protest. A print-out is real evidence, and is to be received as prima facie evidence of the entries. The defence is attempting to add to the costs in this case by making an unreasonable request.'

Judge Nuri Efendi interjected. 'Mr Ayarcı, please elaborate your point.'

Mr Hayri İrdal sat down. Mr Halit Ayarcı stood up.

'My submission is that the technical literature clearly demonstrates that all spreadsheets have significant error rates, and it is our contention that it is obvious that there must be some errors in the documents that affect the figures that my learned friend wishes to have admitted. Indeed, as I have made it clear to my learned friend, the collapse of the banking system in Jamaica in the late 1990s was partly due to the use of spreadsheets and the failure to manage and control them. On this issue alone, I submit that it cannot be right to admit these documents under the bankers' books exception without the electronic versions of the files being subject to analysis by appropriately qualified digital evidence professionals.'

Mr Halit Ayarcı sat down. Mr Hayri İrdal stood up.

'Your honour, as my learned friend is only too well aware, the evidence also benefits from the presumption that mechanical instruments were in order at the material time – a presumption which, I do not need to remind your honour, intentionally included computers. I most strongly resist this potentially expensive and unnecessary challenge regarding the authenticity of the spreadsheets on the basis that this presumption applies.'

Mr Hayri İrdal sat down.

Judge Nuri Efendi considered the submission. 'Mr Ayarcı, notwithstanding the legislative provisions governing business records, the presumption of equipment being properly constructed and operating correctly must be strong, and it is a particularly strong presumption in the case of equipment within the control of the party. Please address this particular issue.'

Mr Halit Ayarcı stood up.

'I appreciate the nature of the presumption, your honour. The exception permits records to be adduced because, in the past, employees entered information into physical books by hand, and this meant they could be relied upon as a record made at that point in time, and one could ascertain at a glance whether somebody tried to change the entries. The justification was that such records were more reliable than

the memory of a witness. This might have been so, but records in electronic form are notorious for being inaccurate for a variety of reasons, and it must be common sense that this rule cannot be relied upon in the twenty-first century.

Let me ask my learned friend what he means that computers are reliable. For instance:

Does my learned friend mean that the spreadsheets are authentic, in that they are the right ones, and they have not been tampered with?

Does he mean that the spreadsheets are valid, in that they contain the information that is claimed of them?

Perhaps he means that the spreadsheets are internally valid, in that the spreadsheets work? If this is the case, what evidence is there that the users of the spreadsheet application checked that the algorithms were correct? My learned friend might also like to confirm if the presumption that computers are reliable includes the maintenance of the spreadsheets and who wrote them, and what qualifications they had to be able to program 'reliably'.

But perhaps he means that the software code of the operating system is reliable? How does he know? How many updates have there been since the spreadsheets began to operate? Were all updates applied? When updates occurred, how did they affect the application software? What is his measure of reliability?

Does he mean that there are no errors of logic that can lead to an incorrect result? What evidence does he have of this, taking into account the number of software code updates to the spreadsheets? Perhaps my learned friend can kindly indicate the number and purpose of each software update since its inception.

Perhaps he means that the employees that input the figures are always accurate? And I presume the system is so reliable that inaccurate inputs are recognized and corrected, and that these corrections are recorded?

No doubt my learned friend can also confirm, because the spreadsheet programs are deemed to be reliable, that there are no errors of omission where the formula is wrong because one or more of its input cells is blank or otherwise incorrect such as referring to the wrong cells?

I ask my learned friend, which part of this process is reliable? All of it? Parts of it? If part of it, which part and for what reason?

But let me finish with another question on the basis that your honour is against my request for electronic versions of the spreadsheets – perhaps my learned friend can assure the court, if only paper versions of the record are to be admitted, that the full information will be provided. That is, he will provide the respective algorithms that undertake the calculations – after all, one does not admit the body of a motor vehicle on its own into evidence to demonstrate the cause of a collision where it is claimed that the brakes failed – one needs to know how the brakes worked and to view the evidence of the braking system. But that is exactly what my learned friend is asking the court to admit: the unsupported assertions of the truth of the contents of spreadsheet programs in the absence of the mechanism by which the data was created.

Finally, before my learned friend responds, we have to consider the requirement that the book is in the custody or control of the bank. This is a significant issue, because, as we now understand it, the spreadsheets in question are maintained in the cloud ……'

Preface

Stephen

The idea for this book came from Helen Vaux (as she then was), the commissioning editor for Butterworths, who sent me an email on 28 January 2004 at 14:27, asking if there was scope for a text covering the discovery, production and admission of electronic information as evidence. (Incidentally, I no longer have this email in electronic form. I only have a version printed on paper with my manuscript notes added on the paper print-out). I thought a book of this nature would be a good idea. The request was for a book of at least 100,000 words. I was not sure that the topic would be sufficient for the length requested (how wrong I was), which is why I suggested that we include individual chapters from a number of common law jurisdictions. Including other jurisdictions was also relevant in my view, because evidence in electronic form knows no physical boundaries. This is how the book developed.

However, as one edition followed another, so the size of the text increased (the first edition comprised 551 pages, the second 812 pages, and the third edition 934 pages). A further increase in size was inevitable with the fourth edition. Unfortunately, and understandably, no publisher wanted to contemplate the publication of a book that would probably run into two volumes. This meant I was placed in a dilemma. I take the view that it is important for lawyers and judges to understand what other lawyers are thinking, and how judges decide cases across the globe on the same topic. Naturally, a decision in one jurisdiction will not necessarily be followed in another jurisdiction for a variety of reasons, but lawyers and judges might wish to be made aware of other decisions that are made given a similar set of facts. This argument aside, it was increasingly obvious that the text could not continue in its previous form.

This was sad, but possibly inevitable.

The book had to be reduced in scope, so I concluded that the only alternative was to structure it around the basic issues facing all judges and lawyers when dealing with electronic evidence, and to base the text on the law of England & Wales, with the usual references to important case law from across the world. It was not an easy decision, because the authors of the various chapters had contributed time and energy into the beginnings of a potentially significant international text. As part of the revision, the topic of electronic disclosure in civil and criminal proceedings has been removed. Although electronic disclosure has always featured in the book, it was inevitable that the subject would eventually merit a separate book, although there is no text, other than my own, *Electronic Disclosure: A Casebook for Civil and Criminal Practitioners* (PP Publishing 2015) – and rapidly dating, that covers electronic disclosure in criminal proceedings.

Stephen and Daniel

Our aim with the revised text is to provide an accurate guide to the state of the law and the technology. Although the focus is on the law of England & Wales, we recognize that a

great deal of important case law and legislation in other jurisdictions is relevant to the issues discussed, and for that reason the text includes references to other jurisdictions when appropriate.

We also recognise that the topic remains in flux, and are in no doubt that the text will continue to evolve, and trust that the electronic nature of this text will facilitate that evolution.

Stephen Mason
Langford, Bedfordshire
stephenmason@stephenmason.co.uk
March 2017

Daniel Seng
Singapore
lawsengd@nus.edu.sg
March 2017

Acknowledgments

Stephen

It is with some sadness that the book has changed so much for the fourth edition. A substantial number of people that have taken part in the text since the first edition in 2007. I wish to thank all of those that have kindly taken part in this book, but who no longer are part of the project. We have had some interesting discussions over the years, and each of those listed below has enriched my knowledge of the topic. Thank you. They are, in alphabetical order of the country in which they are practicing or teaching:

> Robert J. Currie, Professor of Law and Director of the Law & Technology Institute at the Schulich School of Law at Dalhousie University, and Steve Coughlan, Professor or Law at the Schulich School of Law at Dalhousie University and Associate Director of the Law & Technology Institute (Canada)
>
> Ronald Yu, JD, LLM, part-time Lecturer, Faculty of Law, University of Hong Kong and Director and General Counsel of Gilkron Limited (Hong Kong Special Administrative Region, People's Republic of China)
>
> Clive Freedman, Barrister (England & Wales)
>
> Manisha T. Karia, Advocate-on-Record in the Supreme Court of India, New Delhi and Tejas D. Karia, Partner with Amarchand & Mangaldas & Suresh A. Shroff & Co, Advocates & Solicitors in New Delhi (India)
>
> Ruth Cannon, LLB (Dub) BCL (Oxon) BL, Barrister and member of the School of Languages, Law and Social Sciences, Dublin Institute of Technology, and Katie Dawson, BCL Barrister-at-Law (Kings Inns) (Ireland)
>
> Iain G. Mitchell, QC (Scotland)
>
> Associate Professor Daniel Seng and Bryan Tan, Partner at Pinsent Masons LLP (Singapore)
>
> Julien Hofman, Emeritus Associate Professor in the Department of Commercial Law, University of Cape Town and Justin de Jager, BA, LLB, LLM, PGDip, a practicing attorney of the High Court of South Africa (South Africa)
>
> Dr Joseph J. Schwerha IV, M.S., J.D., Associate Professor of Business Law and Technology, Department of Business and Economics, California University of Pennsylvania and John W. Bagby, Professor of Information Sciences and Technology, Pennsylvania State University (United States of America)
>
> Professor Damian Schofield, Director of Human Computer Interaction, State University of New York (Oswego) and Visiting Associate Professor of Digital Forensics, Edith Cowan University, Perth

The text continues with regular and new contributors, mentioned below.

I appreciate Associate Professor Daniel Seng, National University of Singapore, for agreeing to join me in the chapter on the foundation of electronic evidence and for joining me as a joint editor. Daniel has given up a great deal of his time not only to act as a second editor, but also to suggest valuable revisions throughout the text. I am indebted to him for his willingness to continue to take an active part in this project. Daniel has a superb grasp of this subject, which is essential in ensuring that the content can be relied upon by all those taking part in the justice system.

I thank the Institute of Advanced Legal Studies for renewing my Associate Research Fellowship for the academic year 2015–16, which provided me with unlimited use of the IALS Library and Information Services, and thank the librarians at the Institute of Advanced Legal Studies, Gray's Inn, Middle Temple and the Institution of Engineering and Technology Library, Savoy Place, London for their help when responding to requests and offering help.

The previous chapter 5, 'Mechanical instruments: the presumption of being in order', now renamed 'The presumption that computers are "reliable"', has been expanded and revised. It has had the benefit of further review by each of the following:

Dr Chris Elliott, FREng, System Engineer and Barrister

Dr Michael Ellims

Dr David Jackson, CEng MIET FBCS, Global Technical Director, Altran Technologies

Professor Peter Bernard Ladkin, University of Bielefeld; Causalis Limited and Caualis Ingenieur-GmBH

Derek Partridge, Professor Emeritus, University of Exeter

Lorenzo Strigini, Professor of Systems Engineering, City University

Harold Thimbleby, Professor of Computer Science, Swansea University

Martyn Thomas, CBE, FREng, Livery Company Professor of Information Technology, Gresham College; Director, UK Health and Safety Executive

I appreciate the time that each of those mentioned above has given to consider and improve the text. The discussion is vastly improved because of their comments, observations and corrections, and I have shamelessly adopted the vast majority of their recommendations. Nevertheless, the faults remain that of the author.

Also to Harold and Daniel for commenting on the vignette. I began with 'The abacus' for the first edition, and considered stopping for the second edition, but Daniel persuaded me that they were fun, so I have continued with them.

As ever, I remain grateful to my wife, Penelope, for this indulgence.

Daniel

I wish to thank Stephen for his generosity and kindness in granting me the extraordinary privilege to contribute in an editorial capacity to his text on electronic evidence, a culmination of his thought leadership in the area. His boundless energy, intellectual acuity, receptiveness to new ideas, and statesmanship in co-opting contributors and reviewers, make it a real pleasure to work with him. Having been involved since the first edition as a contributor to the Singapore chapter, it is refreshing to return to a close and careful review of the law of evidence in England & Wales, both as a Rupert Cross scholar and an eternally indebted student of Professor Colin Tapper of Oxford University. Indeed, as one of the first scholars in the Commonwealth to conduct a systematic examination of the issues of electronic evidence, Professor Tapper's concise scholarship and legal precision have never ceased to guide and inspire me.

I also wish to thank my Dean, Professor Simon Chesterman, of the Faculty of Law, National University of Singapore, for his thoughtful appreciation, unstinting support and tremendous encouragement for my work, and in giving me the space and time to manage it, in between my teaching and administrative commitments.

Finally, I am obliged to my wife, my mother-in-law and my daughters, for their ceaseless indulgence. I thank my wife, Xu Le, for always being here, and there, for me, and for always being my companion, my confidant, and my best friend.

Joint

Of those that continue with the book, we thank Dr Chris Gallavin, Professor of Law and Deputy pro-Chancellor, Massey University, Palmerston North for taking up the lead in the chapter on hearsay; Burkhard Schafer, Professor of Computational Legal Theory, School of Law, University of Edinburgh for continuing with the chapter on the characteristics of electronic evidence, and Dr George R. S. Weir, Computer and Information Sciences, University of Strathclyde, Glasgow for staying with the introductory chapter. We appreciated Alisdair Gillespie, Professor of Law and Head of Department, Centre for International Law and Human Rights, Lancaster University Law School joining the chapter on encrypted data and Dr Allison Stanfield, for agreeing to join the chapter on authentication. Andrew Sheldon, MSc, of Evidence Talks Limited agreed to remain for the chapter on proof, and Hein Dries, LLM of Viglio Consult also agreed to join this chapter. Hein is a lawyer and a digital evidence professional, and brings a duality to the chapter which is most appreciated. Sandip Patel, QC intended to help out with the chapter on proof, but the pressure on his time meant it was not possible to take advantage of his experience and knowledge. We hope he will be able to become more involved in the future.

All of the authors have taken time out of their already busy commitments to continue to work on this text, and we thank them sincerely for this.

With thanks to all those at the Institute that have helped to produce this fourth edition, in particular Steve Whittle, the IALS Information Systems Manager and Emily Morrell, Head of Publications, School of Advanced Study and Senate House Library, University of London.

We also thank James Kwong and Ashwini Natesan for volunteering to be research assistants on this project and for the excellent work they have accomplished. We wish them well for the future.

Table of statutes

Australia

Canada

England and Wales

New Zealand

United States of America

Table of Statutory Instruments

England and Wales

Northern Ireland

Table of European Legislation

Directives

Regulations

Table of International Legislation

Table of Other Enactments

England and Wales

United States of America

Table of cases

Antigua and Barbuda

Australia

Federal

Industrial Relations Court of Australia

New South Wales

Northern Territory

South Australia

Canada

England and Wales

European Court of Human Rights

European Patent Office Technical Board of Appeal

Fiji

Hong Kong

India

Intellectual Property Office

International Court of Justice

Ireland

New Zealand

Northern Ireland

Patents County Court

Permanent Court of Arbitration

Scotland

Singapore

Tonga

United States of America

Federal

1

The sources of electronic evidence

George R. S. Weir and Stephen Mason

1.1 Various devices are capable of creating and storing data in digital form, and such data may serve as evidence. The aim of this chapter is to introduce the reader to the technologies, their underlying principles and the general characteristics that set evidence in digital form apart from evidence in analogue or physical form. The content of this chapter does not deal with any of these matters in depth. Neither does it aim to be a comprehensive review of the devices and technologies that create electronic evidence. Rather, the aim is to provide a broad brush introduction to the relevant technical issues, and to highlight features that a digital evidence professional and a legal professional should be concerned about when investigating electronic evidence and dealing with electronic evidence issues.

Digital devices

1.2 Historically, the term 'computer' was often used to describe almost any form of processing unit. Now, digital computation and storage facilities are characteristic of many devices that seem far removed in form and function from traditional computers. Such devices include games consoles, wearable technologies (e.g., fitness trackers, smart watches) and 'smart' domestic components (e.g., smart energy meters, automated central heating systems). Most of these digital devices share important features with more recognizably conventional computing devices such as desktop computers, laptops and computer tablets. These features are based on what is sometimes called an input-processing-output model:

> The device receives an *input* of some sort, by way of a local file, sensor, mouse, keyboard or through a communication channel (such as a network connection).
>
> It *processes* the information.
>
> It produces an *output* to a display, local file or printer, for instance.
>
> It must be able to *store* (and/or relay) information.
>
> It must be able to *control* what it does.

1.3 In the following, we detail the role played by the main components in digital processing systems (digital devices).

The processor

1.4 The processor, also called the central processing unit (CPU), is the functional core constituent of every such device, and is itself made up of a number of constituent parts. Together, these parts receive data, perform logical or arithmetic operations and output the results. The results are passed to a local storage facility or a display unit, or 'uplinked' via a network connection to another device.

George R S Weir and Stephen Mason, 'The sources of electronic evidence', in Stephen Mason and Daniel Seng (eds.), *Electronic Evidence* (4th edn, University of London 2017) 1–17.

Software

1.5 Software consists of programs that give instructions to the digital device. There are two main categories of software: system software and application software.

System software

1.6 As the name suggests, system software is required for the basic operation of a device. The set of software programs that manage the basic operation of a computer is called the *operating system*. The operating system controls the flow of data, allocates memory, and manages any hardware components of the device, such as the display, input device(s), network interaction, etc. The operating system also permits the user to manage any user-specific files, enabling multiple users to share the use of a computer, and acts as an interface between the hardware and the application software.

Application software

1.7 Broadly speaking, for more traditional computing devices such as desktop computers, laptops and tablets, the application software provides the user-facing side of the system. This is 'special purpose' software that enables the user to undertake specific kinds of tasks on the computer. These include word processing, desktop publishing, web browsing, email, social networking, preparing and delivering presentations, performing complex sets of numerical calculations and the like. Examples of application software include Microsoft Word, Internet Explorer, Outlook, PowerPoint, Excel and LibreOffice. These and other application programs represent the main reasons for which most people use computers (that is, to perform specific tasks, made simpler by means of the computer and its application software). For other digital devices, the user may only engage the application software through a limited range of functions, such as status checks on a fitness tracker or energy consumption from a smart meter.

The clock

1.8 One further component must be discussed in relation to the operation of digital devices: the clock. The clock serves two functions:

> (i) It is a device that produces pulses of time to ensure that events are synchronized and occur in a predictable order. The clock coordinates all the components of the CPU. Each step in any operation must follow in sequence, and some operations run at different speeds. System operations are synchronized to the pulses of the electronic clock. The frequency of pulses is controlled by a *phase locked loop* (PLL), which, in turn, is regulated by a quartz crystal. The speed at which the crystal oscillates, the step-up ratio of the PLL, and the number of steps that each instruction requires, will determine the speed at which the computer operates.
>
> (ii) The clock also often serves to keep the time of day and date in a human sense. Larger computer systems synchronize their clocks with a reliable time source available over the Internet, using a system interface such as the Network Time Protocol. This allows devices attached to the Internet to synchronize their time settings (taking into account geographical locations and time zones). There are two important reasons to provide for the synchronization of time. The first purpose is to ensure that events occur on time, and in the correct sequence. This permits events to be scheduled and enables the fact that they have occurred to

> be registered accurately. The second purpose is to enable a person to retrieve information concerning past events, including establishing when the events occurred and the sequence in which they occurred. This is only possible if accurate time stamps are available. Examples include the time stamping mechanism for the purpose of authentication, digital signatures and the diagnosis of faults recorded on system event logs.

1.9 In most implementations, the built-in clock is powered by a battery and runs continuously even when the device is switched off. Devices that have lain for a long time without being powered on may not 'boot up' when they are turned on, because the battery has run down and may require recharging or replacing. We should also note that the clock in digital devices is often imprecise (like an inexpensive wristwatch). Usually, the clock can be adjusted (and even incorrectly set) manually. This can result in the system clock being slightly incorrect (through 'drift' in time keeping) relative to the actual time in the local region. Such inaccuracy may affect both uses of the clock indicated above, i.e., event scheduling and logging, since both aspects may depend upon the time as derived from the system clock. Where time accuracy is important, the clock usually requires occasional adjustment to bring the time back into line with better reference sources (such as Internet time servers). This is a matter of some significance, since unquestioned and out of context assumptions about the accuracy or otherwise of a clock may result in a misleading conclusion.

Time stamps

1.10 From the perspective of electronic evidence, the system clock often plays a vital role in time stamping events. For instance, the operating system uses the date and time settings to annotate its record of events such as the creation or modification of a file. In computers, such information is often referred to as file 'metadata' (the data that describes or interprets the base data), since the date and time information is associated with the file, but is not part of the data in the file or data that the user has any direct control over. Time stamps are also recorded against system events such as user logins, password changes and, depending upon the purpose of the device, sensor-recorded events such as number of steps walked by the wearer. The time and date information associated with such events is recorded in system log files (event logs). Such logs are often an important source of event sequence information and afford insights on purported specific user activity.

1.11 As noted earlier, the clock in a computer can be set by the user and may not be configured to maintain the correct current time (such as using the Network Time Protocol). Incorrect time settings will be reflected in the date and time stamps subsequently recorded by the system. Obviously, this potential anomaly must be considered when dealing with data that is time stamped. Since the time zone is also set in the system, an incorrect choice of zones may result in an incorrect current date or time. In addition, because of the critical role the clock plays, it features a great deal in electronic evidence, particularly where it is manipulated by the defendant to hide evidence of changes made to critical evidence.[1]

1 Chet Hosmer, 'Proving the integrity of digital evidence with time' (2002) 1 Intl J of Digital Evidence 1; Chris Boyd and Pete Forster, 'Time and date issues in forensic computing – a case study' (2004) 1 Digital Investigation 18; Malcolm W Stevens, 'Unification of relative time frames for digital forensics' (2004) 1 Digital Investigation 225.

Memory and storage

1.12 In order to retain programs, output results and other data on which programs operate, digital devices rely on storage. There are generally speaking two forms of storage: primary storage and secondary storage. Primary storage is storage that is directly accessible by the processor. It typically takes the form of semiconductor memory such as:

> (i) An internal storage chip known as *random-access memory* (RAM).[1] This chip is capable of repeatedly storing (writing) and retrieving stored data (reading).
>
> (ii) An internal storage chip that is capable of storing data once, but does not allow the data to be re-written. Once data has been entered, this type of chip only allows the data to be read. This is called read-only memory (ROM).[1]
>
> (iii) An internal storage chip that stores data and behaves as a ROM during its normal operation, but permits data to be erased and replaced. This form of device is known as *erasable programmable read-only memory* (EPROM).[2] A flash ROM is a type of EPROM.

1 'Random-access memory' (*Wikipedia*) <https://en.wikipedia.org/wiki/Random-access_memory>.
2 'EPROM' (*Wikipedia*) <https://en.wikipedia.org/wiki/EPROM>.

1.13 Secondary storage is storage that is not directly accessible by the processor. Where data on which it is stored is required, the processor will use its input/output channels to obtain access to secondary storage and transfer the required data into the primary storage. Unlike primary storage, secondary storage is non-volatile: it retains its data when the device is powered down. Hard disk drives (HDDs) and USB 'thumb drives' as storage media are typical forms of secondary storage. They may be permanently attached to the computer (internal storage), or attached when required (external storage). Other forms of external storage may be less proximal to the computer, such as network-attached storage (NAS),[1] tape drives or 'cloud' storage.

1 'Network-attached storage' (*Wikipedia*) <https://en.wikipedia.org/wiki/Network-attached_storage>.

1.14 Because secondary storage is non-volatile, the hard disk and associated offline storage media are a significant source of electronic evidence for a device. But the fact that primary memory is volatile does not mean that its data cannot be retrieved. An experiment on 'freezing' RAM chips before physical removal and transfer to a different computer revealed an unusual context in which RAM data may be recovered from the treated chips.[1]

1 J Alex Halderman and others, 'Lest we remember: cold boot attacks on encryption keys', in *Proceedings of the 17th Conference on Security Symposium* (USENIX Association 2008), and (2009) 52 Communications of the ACM 91, <https://citp.princeton.edu/research/memory/> (abstract only)

Data storage facilities

1.15 The increasingly varied ways of storing digital data and the variety of storage contexts means that locating relevant data as prospective evidence may not be a simple matter. Data may be stored locally to a computing device, such as hard disks, DVDs or CDs, flash drives, memory sticks, or micro memory devices (as commonly found

in smartphones). But data may also be stored remotely such as on network-attached storage, remote networks or 'cloud' facilities. Of concern to many digital investigators is the difficulty inherent in locating and obtaining legal access to data that is stored remotely from an individual's computer. The common data storage contexts are summarised in the table below.

Memory type	Volatile	Local
RAM	Yes	Yes
HDD (internal)	No	Yes
HDD (portable)	No	Perhaps
Flash/USB	No	Perhaps
CD/DVD	No	Perhaps
Network	No	Perhaps
Cloud	No	Typically No

Lost data

1.16 A digital evidence professional may be able to detect a range of 'lost' data on a hard disk or other storage media:

> (i) Where a user intentionally marks portions of the hard disk as 'bad', he can hide substantial amounts of data in those portions that could not be seen without the use of an appropriate disk diagnostic or examination tool (since the operating system will automatically avoid making any use of these 'bad sectors').
>
> (ii) When the user deletes data, it remains on the disk until the old file is overwritten by new data. Only the system's pointers in the filing system are deleted. Even where part of a file has been overwritten, it is often possible to recover part of the deleted file if the entire set of disk blocks containing the original file has not been completely overwritten.

1.17 However, it does not follow that the recovered data is genuine or trustworthy evidence just because it is found. There are numerous contexts in which data may be lost or damaged, and this will affect the credibility of any resulting data that is recovered. Examples include the corruption or loss of data from errors in the program, and interference with the data from extrinsic sources.[1] Further, it should be observed that the reliability of the evidence would also be affected by the way in which a digital evidence professional carries out the examination and recovers the data. If the process of investigation affects the evidence, it will be less reliable.

1 Peter Sommer, 'Downloads, logs and captures: Evidence from cyberspace' (2002) 8 CTLR 33; Eoghan Casey, 'Error, uncertainty, and loss in digital evidence' (2002) 1 Intl J of Digital Evidence; Caroline Allinson, 'Audit trails in evidence – A Queensland case study' (2001) 1 JILT; and 'Audit trails in evidence: Analysis of a Queensland case study' (2003) 2 JILT.

Data formats

1.18 Digital data may be broadly classified into binary data, where the information is represented in binary form, and text data, including alpha, numeric and punctuation data. Text can be entered into the computer by a range of methods:

(i) The typing of letters, numbers and punctuation, mainly when using the keyboard.

(ii) Scanning a page with an image scanner and converting the image into data by using *optical character recognition* (OCR)[1] software.

(iii) Using a *bar code*. The bar code represents alphanumeric data. The bar code is read with an optical device called a wand. The scanned code is converted into binary signals, enabling a bar code translation component to read the data.

(iv) Reading the magnetic stripe on the back of a credit card.

(v) Voice data, where a person speaks into a microphone capable of recording the sounds. This form of data, as well as video data, is encoded in binary form.

(vi) Speech to text. In this instance, the user speaks into a microphone that is connected to the computer and a dedicated software application analyses the input signal and converts this to a textual representation of the spoken words.

1 'Optical character recognition' (*Wikipedia*) <https://en.wikipedia.org/wiki/Optical_character_recognition>.

1.19 To enable a user to view text and numbers, and to see images or hear sound, the binary form of the data must be converted using a code. Binary information can be represented using the binary (base 2) number system, although it is more common to represent computer numbers in octal (base 8) or, most commonly of all, hexadecimal (base 16).

1.20 A range of codes exists for text data. Some of the codes that are in common use are known as Unicode,[1] American Standard Code for Information Exchange (ASCII),[2] Extended Binary Code Decimal Interchange Code (EBCDIC),[3] and Unicode Transformation Format-8 (UTF-8),[4] which is the standard character code used over the Internet that is capable of encoding all possible characters. Most computers now use Unicode and ASCII. Tools are available to display binary data used in computers to enable a digital investigator to view features that are normally not visible to the computer user. For instance, documents stored in the Microsoft Word format contain application metadata that are normally not visible. By using certain types of software programs, a digital evidence investigator is able to view all aspects of the data and such data may reveal crucial information that may help an investigation.

1 'Unicode' (*Wikipedia*) <https://en.wikipedia.org/wiki/Unicode>; J Klensin and Michael Padlipsky, 'Unicode format for Network Interchange' (2008) RFC 5198 <https://tools.ietf.org/html/rfc5198>

2 'ASCII' (*Wikipedia*) <https://en.wikipedia.org/wiki/ASCII>; Vinton Cerf, 'RFC 20 – 'ASCII format for Network Interchange' (1969) RFC 20 <https://tools.ietf.org/html/rfc20>.

3 'EBCDIC' (*Wikipedia*) <https://en.wikipedia.org/wiki/EBCDIC>; J M Winett, 'The EBCDIC codes and their mapping to ASCII' (1971) RFC 183 <https://tools.ietf.org/html/rfc183>; R T Braden, 'EBCDIC/ASCII mapping for Network RJE' (1972) RFC 338 <https://tools.ietf.org/html/rfc338>.

4 'UTF-8' (*Wikipedia*) <https://en.wikipedia.org/wiki/UTF-8>; F Yergeau, 'UTF-8, a transformation format of ISO 10646' (2003) RFC 3629 <https://tools.ietf.org/html/rfc3629>.

Starting a computer

1.21 Every time a digital device is switched on, various components must interact with each other for it to begin working. This is called the start-up process or 'booting' the system. Most devices have a program in read-only memory called variously a *boot loader*, *boot process*, *boot strap* or *initial program load*. It is this program that enables the system to start. In general terms, this is how it works:

> (i) When the system is powered on, control is first transferred to the basic input and output system (BIOS),[1] a program located permanently in the ROM of the device.
>
> (ii) The BIOS tests the various components of the system, verifying that they are active and working. The results of the various tests it carries out may appear on the system output. The boot process can also clear local primary memory of all historical data and metadata. The BIOS locates the first (or default) secondary storage device, looks for an operating system on the storage device, and passes control to the operating system's boot record (a set of instructions starting at a specific location on the storage device).
>
> (iii) The boot record takes control of the system. This program also contains a boot loader, which, in turn, loads and tests the configuration before loading the operating system.
>
> (iv) Finally, the operating system will display any startup dialogue (for instance, the identity of the mobile telephone service provider), and, if the user is authorized (for instance by providing a code), grant access to application-level programs and the user can take control of the device through the application.

1 'BIOS' (*Wikipedia*) <https://en.wikipedia.org/wiki/BIOS>.

Types of evidence available on a digital device

1.22 A digital evidence professional can make a range of evidence available from a digital device. This section provides an outline of some of the types of evidence that can be gleaned.

Files

1.23 A wide range of application software is used on computers, laptops, tablets and mobile telephones, including programs that enable a user to send messages, prepare spreadsheets, databases and text documents, take digital photographs, and create multimedia and presentations. The files, which will store messages, spreadsheets, databases, texts, photographs, multimedia and presentations, may themselves be electronic evidence. A great deal of data can be retrieved, depending on the method of storage, the media on which it is stored, and the manner in which the device manages data storage.

Imaging

1.24 Any digital forensic investigation will begin by 'imaging' the device on which electronic evidence may reside. The imaging process is a non-destructive process that creates an exact external digital copy of any data on the device. Subsequently, all data

investigation should be performed on the imaged copy and not on data stored on the original device.

System and program logs

1.25 In most modern operating systems such as Windows and Linux, virtually anything and everything happening on and to the system is recorded in the form of logs in some manner. This includes information about system events, including the startup of applications and various classes of error messages. Information in the logs may help to determine, for instance, how an unauthorized computer user obtained access to a system with the intent of stealing information from the computer. It may also be possible to configure the systems log (syslog) such that the log messages can be sent to another networked system while retaining a local copy. As a result, if a hacker acquires root privileges on a networked UNIX system, for instance, and wants to erase something from the local logs, he would not be able to erase the datum from the remote logs to remove all traces of his intrusion unless he also has the appropriate privileges on the remote machine.

1.26 Unlike UNIX systems, the Windows operating system also includes a 'registry'. This is a store of data that contains a great deal of information, including a comprehensive database containing information on every program that is compatible with Windows that has been installed on the computer. It also includes information about the purported user of the computer, the preferences exercised by the user, information about the hardware components, and information about the network (if it is connected to a network). The values stored in the registry are in hexadecimal format, but can be converted to ASCII. An example of the type of information that the registry can provide to an investigator is the AutoComplete data for a user of Internet Explorer visiting a particular website such as his name, address, telephone number, email address and passwords. In addition, it is possible to establish when the user last downloaded a file from the Internet, and the first page the user visited from the registry.[1]

1 Although it does not follow that a user clicked on a website address that has been recorded in a temporary cache file, for which see the case of *State of Connecticut v Julie Amero* (Docket number CR-04-93292; Superior Court, New London Judicial District at Norwich, GA 21; 3, 4 and 5 January 2007). For an exhaustive analysis of this case, see Stephen Mason (ed), *International Electronic Evidence* (British Institute of International and Comparative Law 2008) xxxvi–lxxv.

Temporary files and cache files

1.27 When a computer connects to the Internet, a range of information about its activities is recorded and retained locally, including the websites that have been visited, the contents that were viewed and any newsgroups that were visited. For the purpose of enabling the browser to improve the user experience and speed up browsing, temporary copies of websites that have been visited are stored in cache folders. These folders contain fragments of the web page, including images and text. Some versions of software will retain in more than one local file location information about the websites visited.

1.28 It is important to understand the legal consequences of the temporary files and cache files. This is exemplified in the case of *Atkins v Director of Public Prosecutions*.[1]

In this case, Dr Atkins, a university lecturer at the University of Bristol, Department of English, had browsed the Internet for indecent photographs of children. He deliberately saved a number of such photographs as files in the J directory, but he did not know that these photographs were also cached in the temporary cache folder of Internet Explorer. He was convicted on ten offences of having in his possession indecent photographs of children, nine in the form of temporary files in the cache folder and one from the J directory. He was acquitted of a further 24 charges, some of which related to the files deliberately saved in the J directory. Both his and the prosecutor's appeals were allowed. Simon Brown LJ and Blofeld J held that Dr Atkins should not have been convicted of possession in respect to the photographs stored in the cache, because he was not aware of its existence or what it did, and therefore could not be said to have knowingly had possession of these particular photographs. He should only have been convicted of intentionally placing the photographs in the J directory, because he knew what he was doing. The court ordered that the case be remitted with a direction to convict Dr Atkins of the offences where he deliberately saved photographs in the J directory.[2]

1 *Atkins v Director of Public Prosecutions*; *Director of Public Prosecutions v Atkins* [2000] 1 WLR 1427 (QB); for a US case based on similar facts with an identical outcome, see *United States v Kuchinski* 469 F.3d 853 (9th Cir. 2006).

2 In *Clifford v Chief Constable of the Hertfordshire Constabulary* [2011] EWHC 815 (QB), Mr Justice Mackay observed that the prosecution were fully aware of this issue, but prosecuted Mr Clifford in any event: a prosecution that was eventually determined to be malicious; see also *Clifford v Chief Constable of the Hertfordshire Constabulary* [2008] EWHC 3154 (QB) and *Clifford v Chief Constable of the Hertfordshire Constabulary* [2009] EWCA Civ 1259.

1.29 In addition to browser caches, Windows and UNIX systems also have *paging file* or *swap space*. This is an area of disk that is used as *virtual memory*. In the event that the applications being run on the system require more RAM than a system has available, low priority applications that are running are copied to the virtual memory and the RAM they are using freed for use by applications with a higher priority. Swap space is rarely cleaned during the normal operation of the system. This means that when a system needs to be forensically analysed, it is often the case that useful data associated with applications, which may not even be running at the time, can be found by analysing the content of the swap space. This can also apply to data that is normally stored on the standard file system in an encrypted form. Depending on the application and the precise circumstances, some applications may allow unencrypted copies of the data to be stored in the swap file.

Deleted files

1.30 File systems keep a record of where data are located on a storage medium. The way data are stored will differ, depending on the software and the architecture of the method used to allocate blocks of storage for files (the file system architecture). In simple terms, the location of data on a storage medium is controlled by a file system. For instance, the storage medium can be divided into partitions, and where this is the case, the file will be stored on a particular location in a partition. When a file is deleted, the instruction to delete removes the pointer to the location of the file, but does not actually delete the file. For this reason, in the vast majority of cases, it is possible to recover data that have been deleted, depending on the amount of disk

writing activity that has been performed between the deletion of the file and the forensic analysis.[1] Alternatively, when a device that is claimed to be non-functional is forensically restored or unlocked, it may be possible to discover or infer evidence of wrongdoing on the device. This is illustrated by the case of *Sectrack NV v Satamatics Ltd*[2] concerning the misuse of confidential information. One of the defendants was in possession of a Blackberry device, which he claimed was frozen or locked. When the device was 'unlocked', it automatically downloaded various emails that the defendant received, which implicated him in the misuse of confidential information.[3] Since this case, manufacturers of hand held devices have developed extensive back-up systems that permit the back-up of device data to other devices and storage facilities. In future, without the use of encryption, it will be relatively difficult to delete data sufficiently for it to be beyond recovery.

1 Andy Jones and Christopher Meyler, 'What evidence is left after disk cleaners?' (2004) 1 Digital Investigation 183.
2 [2007] EWHC 3003 (Comm).
3 [2007] EWHC 3003 (Comm), [7].

Mobile devices

1.31 Hand held devices are now ubiquitous. These include the use of tablets and smartphones that combine personal computer functionality with telephone and camera. Such devices are computers, since they have a CPU, memory, keypad or mouthpiece (input) and a screen or earpiece (output). And like computers, hand held devices have ROM and RAM. The ROM stores the operating system and any essential software required for the device to function. The RAM is used to store other software and data that the user may wish to retain. More recently, these devices are equipped with a programmable ROM known as flash-ROM, a form of solid-state memory chip that is capable of retaining content without power.

1.32 Other types of specialist mobile devices include digital music players and ebook readers that can use wireless technology to download large volumes of data from a main computer. All these devices, together with laptop computers, are increasingly used by organizations as components in an extended information technology infrastructure. Where relevant, such devices may be investigated for electronic evidence, although the amount of information that can be obtained will vary. For instance, while one may only find a list of the most recent telephone numbers called from an ordinary mobile telephone, a smartphone will probably yield substantial amounts of data, including emails and other data from a network that might aid an investigation.

1.33 The examples given above emphasize the types of electronic evidence that can be revealed by means of a forensic examination, including hidden or deleted data. Only a highly skilled person could remove all traces of evidence on a computer, and such skills are very rare. Some forensic techniques exist that can recover data even when it has been strictly overwritten on disk. Whether these techniques will be used or deployed will of course depend on the type and value of the data sought to be recovered.

Networks

1.34 Gone are the days when most computers stood alone on a desk. The majority of computers are now connected, or are intermittently connected, to other computers, or a network. Given the trails left by the assortment of logs and files in computers, going online can produce electronic evidence in abundance, including the using of email, connecting to the Internet and viewing websites, and transferring of files between computers. Other sources of electronic evidence can be obtained from server logs, the contents of devices connected to the network, and the records of traffic activity. In many instances, it could be that the only evidence that will be available is evidence on a network, because the perpetrator of a crime may have successfully persuaded the victim to destroy evidence by disposing of his hard drive and any other hardware.

Types of network

Internet

1.35 The development of the Internet was brought about because the military in the United States of America recognized the need to ensure military communication networks could continue to communicate, even if important parts of the infrastructure were damaged beyond repair. Since the introduction of the World Wide Web, it has become easier for people to use the Internet. Other networks also exist that operate at higher speeds, such as the Internet2. When a computer connects to the Internet, it uses a set of protocols called *Transmission Control Protocol/Internet Protocol* (TCP/IP).[1] This set of communication standards can be regarded as a common language that enables various types of networks to communicate, each with the other. When a computer is connected to a network, it is referred to as a 'host'. The computer uses a *modem* or a *network interface card* (NIC)[2] to send and receive information, although medium-sized and large organizations will have a *Local Area Network* (LAN)[3] gateway to the Internet. A computer, or host, that is connected to two or more networks is called a 'router' if it mediates the passage of traffic between them, and if the networks have different addresses. Most networks use bespoke routers. Routers are a very important part of a network, because they act to direct data from one network to another, filter traffic that is not permitted, and keep logs of activity. Most routers maintain system logs, which may vary in terms of the quantity of data and the amount of detail in each log entry.

1 'Internet protocol suite' (*Wikipedia*) <https://en.wikipedia.org/wiki/Internet_protocol_suite>; Vinton Cerf, 'Specification of Internet Transmission Control' (1974) RFC 675 <https://tools.ietf.org/html/rfc675>; F Baker, 'Requirements for IP Version 4 routers'(1995) RFC 1812 <https://tools.ietf.org/html/rfc1812>.

2 'Network interface controller' (*Wikipedia*) <https://en.wikipedia.org/wiki/Network_interface_controller>.

3 'Local area network' (*Wikipedia*) <https://en.wikipedia.org/wiki/Local_area_network>.

1.36 A further component of the modern communication infrastructure is the *server*, often viewed as a very powerful computer that provides a range of clients with a service, for instance, hosting an organization's web service or email facility. Some servers, such as web servers, permit anyone to obtain access to its resources without limitation. Other servers, such as email servers, only permit authorized users to obtain access to the service, usually by means of a username and password. Sources

of electronic evidence from servers include logs recording when a user connects to a server, whether to grant access to the Internet or whether to download email.

Corporate intranets

1.37 An intranet, usually run by a large organization, is a private network that in principle is only available to members and employees of the organization or others with authorization to obtain access to and use the information contained on the network. The intranet may look like a smaller version of the Internet, providing websites, mail servers and time servers amongst other facilities. Usually situated within the corporate firewall, an intranet is built to support the internal needs of the organization, and to improve workforce connectivity and business operations. As such, it generally aims to keep those outside the organization from gaining access, and is usually well protected.

Wireless networking

1.38 A further development in this form of networking is wireless technology. One implementation of wireless networking is Wi-Fi[1] (a mark used by the Wi-Fi Alliance), mainly through the 2.4 GHz and 5 GHz radio bands based on the 802.11 communications standard.[2] Another wireless technology, known as Bluetooth,[3] is a wireless technology standard for exchanging data between devices over short distances using ultra high frequency (UHF) radio waves. From an evidential perspective, logs exist to record the use of wireless networks, affording evidence of the use that a device has made of a network.

1 'Wi-Fi' (*Wikipedia*) <https://en.wikipedia.org/wiki/Wi-Fi>.

2 The number 802 is the name given to the interoperability standard developed by the Institute of Electrical and Electronic Engineers for Local Area Networks and Metropolitan Area Networks, and Wi-Fi is based on 802.11, which is a sub-set of the 802 standard relating to wireless local area networks.

3 'Bluetooth' (*Wikipedia*) <https://en.wikipedia.org/wiki/Bluetooth>.

Cellular networks

1.39 A cellular network or mobile network is a communications network that enables portable devices such as cellular telephones to communicate with each other. The network is made up of a number of cell sites (base stations) within a defined geographical area. An individual connected to a cell site can make and receive calls over the network. Each cell site is connected to a central computing infrastructure, comprising telephone exchanges or switches, which are in turn connected to the public telephone network. This infrastructure processes the calls by routing them to their destination, and retains logs for the purpose of sending out bills, maintenance and, if necessary, carrying out investigations. The most recent developments in the cellular technology include *General Packet Radio Services* (GPRS),[1] *Third Generation* (3G)[2] *Universal Mobile Telecommunications System* (UMTS),[3] and the *Fourth Generation* (4G)[4] *Long Term Evolution* (LTE)[5] standard, developments that provide for faster transmission rates and enable applications such as mobile web access, IP telephony, gaming services, high-definition mobile TV, and video conferencing. These supplant and will eventually replace the *Global System for Mobile Communications* (GSM)[6] standard, which, while incorporating encryption mechanisms, is now considered to have security flaws which are complex, though feasible, to exploit.

1 'General Packet Radio Service' (*Wikipedia*) <https://en.wikipedia.org/wiki/General_Packet_Radio_Service>.
2 '3G' (*Wikipedia*) <https://en.wikipedia.org/wiki/3G>.
3 'UMTS (telecommunication)' (*Wikipedia*) <https://en.wikipedia.org/wiki/UMTS_(telecommunication)>.
4 '4G' (*Wikipedia*) <https://en.wikipedia.org/wiki/4G>.
5 LTE (telecommunication) (*Wikipedia*) <https://en.wikipedia.org/wiki/LTE_(telecommunication)>.
6 'GSM' (*Wikipedia*) <https://en.wikipedia.org/wiki/GSM>; H Haverinen and J Salowey (eds.), 'Extensible Authentication Protocol Method for Global System for Mobile Communications (GSM) Subscriber Identity Modules (EAP-SIM)' (2006) RFC 4186 <https://www.ietf.org/rfc/rfc4186.txt>.

1.40 A mobile telephone has several numbers that identify the device. The manufacturer includes an *Electronic Serial Number* (ESN)[1] or the *International Mobile Equipment Identity* (IMEI)[2] number as a code to uniquely identify mobile devices. The *International Mobile Subscriber Identity* (IMSI)[3] number is a unique identification number, typically provisioned in the SIM card of the telephone to identify the subscriber of a cellular network. To prevent the subscriber from being identified, this number is rarely sent. What is sent in its place is the *Temporary Mobile Subscriber Identity* (TMSI),[4] which is randomly generated and assigned to the telephone the moment it is switched on, to enable the communications between the mobile device and the base station. Finally, the *Mobile Identification Number* (MIN) or *Mobile Subscription Identification Number* (MSIN)[5] is the unique telephone directory number for that mobile subscription that is used to identify a telephone. It is derived from the last part of the IMSI.

1 'Electronic serial number' (*Wikipedia*) <https://en.wikipedia.org/wiki/Electronic_serial_number>.
2 'International Mobile Station Equipment Identity' (*Wikipedia*) <https://en.wikipedia.org/wiki/International_Mobile_Station_Equipment_Identity>.
3 'International mobile subscriber identity' (*Wikipedia*) <https://en.wikipedia.org/wiki/International_mobile_subscriber_identity>.
4 'Mobility management' (*Wikipedia*) <https://en.wikipedia.org/wiki/Mobility_management#TMSI>.
5 'Mobile identification number' (*Wikipedia*) <https://en.wikipedia.org/wiki/Mobile_identification_number>.

1.41 To ensure the telephone company knows the correct base station to direct the call, the position of the telephone is constantly tracked when it is switched on. Thus, there is a broad range of electronic evidence associated with the use of a mobile telephone, including where the telephone was located geographically, details of calls made and received, and the recovery of the contents of text messages.[1] Where a telephone is capable of being used in other ways, such as making micro-payments, data relating to such services are also capable of being retrieved.[2]

1 In *R v Brooker* [2014] EWCA Crim 1998, Brooker falsely accused her former partner, Paul Fensome, of various crimes, including rape and assault. Cell site analysis determined that Brooker was not at various locations as she claimed. In addition, because Mr Fensome retained all of the text messages exchanged with Brooker, it was possible to establish that the relationship between the two was not as alleged by Brooker.
2 Svein Yngvar Willassen, 'Forensics and the GSM mobile telephone system' (2003) 2 Intl J of Digital Evidence.

Types of network applications

Email

1.42 A significant amount of correspondence undertaken within and between organizations takes the form of the exchange of email. Email is, essentially, an unstructured form of communication, whose content determines its purpose:

> (i) An email discussing official business between employees internally is an internal memorandum.
>
> (ii) A similar email sent out to a third party relating to official business is an external communication, and by being sent with the same corporate information that is contained on the stationery, should be treated as official stationery.
>
> (iii) An extension of a telephone conversation, confirming something, for instance, is a note to be added to a file, whether it is sent to people within the organization or to external addressees, or a mix of internal and external addressees.
>
> (iv) A note to a friend to say you enjoyed the party last night, or to colleagues inviting them to join you in a glass of port and a slice of Dundee cake to celebrate your birthday, is an item of private correspondence using the organization's resources. The use of email for this purpose may or may not be authorized by the organization.

1.43 Email is an important source of electronic evidence. However, emails should be treated with some discretion, because a person can conceal his identity and hide behind a false email address with relative ease. It is very straightforward to send an email that appears to come from someone other than the real source. Forging emails might be effortless, but email is freely admitted into legal proceedings, both criminal and civil.

1.44 To obtain access to email, it is necessary to interact with two different services, one for outgoing mail and one for incoming mail. These services may, or may not, be provided by the same server. To read email, the individual must direct the email program to connect to a mail server using one of a number of protocols, the most common of which are: *Post Office Protocol* (POP),[1] *Internet Message Access Protocol* (IMAP),[2] and a proprietary Microsoft protocol called *Messaging Application Programming Interface* (MAPI).[3]

1 'Post Office Protocol' (*Wikipedia*) <https://en.wikipedia.org/wiki/Post_Office_Protocol>; J Myers and M Rose, 'Post Office Protocol – version 3' (1996) RFC 1939 <https://tools.ietf.org/html/rfc1939>.
2 'Internet Message Access Protocol' (*Wikipedia*) <https://en.wikipedia.org/wiki/Internet_Message_Access_Protocol>; M Crispin, 'Internet Message Access Protocol – Version 4rev1' (2003) RFC 3501 <https://tools.ietf.org/html/rfc3501>.
3 'MAPI' (*Wikipedia*) <https://en.wikipedia.org/wiki/MAPI>.

1.45 The POP protocol (POP3 is the most widely used version) permits the user to read his email by downloading it from a remote server and onto the storage facility of his local computer or device. Once the email has been downloaded from the server, it is automatically deleted from the live server, but probably not from the back-up server that will invariably be used by the mail service provider for the purpose of recovering from a failure for any reason. By contrast, the IMAP protocol (IMAP4 being the most widely used) enables the user to leave all his email on the mail server. Keeping all the email on a single server can be an advantage for an organization because the email for

the entire organization can be backed up from a central location. However, the problem with keeping all email communications on the server is that the server may eventually become overloaded due to the volume of data. Both POP and IMAP protocols require a user to have a username and a password before the user can obtain access to the mail download service. In addition, the protocol servers keep logs of who checked emails and when they were checked. This enables an investigator to look for evidence of email traffic even where a user has deleted all of his emails.

1.46 Outgoing email uses a different protocol called *Simple Mail Transfer Protocol* (SMTP),[1] although MAPI also supports outgoing email. The servers supporting SMTP do not normally require a user to use a password. This makes it very easy for an individual to forge a message. However, the SMTP server may keep a log of the messages that pass through the system.

1 'Simple Mail Transfer Protocol' (*Wikipedia*) <https://en.wikipedia.org/wiki/Simple_Mail_Transfer_Protocol>; J Klensin (ed.), 'Simple Mail Transfer Protocol' (2008) RFC 5321 <https://tools.ietf.org/html/rfc5321>.

1.47 When an email is sent from a computer, it will pass on to one of a number of *Message Transfer Agents* (MTA). The MTAs act in the same way as post offices. A local MTA will receive the email. Upon receipt, it will add to the top of the email message received the current time and date, the name of the MTA, and other additional information. This information in what is called the *header* of the email. As the message passes through various MTAs, each MTA will add further date and time stamps to the header. The most recent information will be at the top of the header.

1.48 Another item of information that tends to be collected in the header is the *Internet Protocol* (IP) address of the computer or system connecting to the server. Technically astute users of email who may wish to hide their identity can send messages through anonymous or pseudonymous re-mailing services. When email is sent through such a re-mailing agent, the header information may be stripped before the message is sent on to its destination. However, some other forms of electronic evidence are transferred during such a process, and it is possible for forensic investigators to attempt to find evidence that may be useful.[1]

1 See Craig Earnshaw and Sandeep Jadav, 'E-mail Tracing' (2004) 15 Computers & Law 7 for an introduction.

Instant messaging

1.49 Instant Messaging (IM) is a form of online communications service that enables the user to transmit a variety of text, voice and image messages with other individuals in real time over the Internet. This form of communication is similar to a conversation over the telephone, but the users typically communicate by typing messages into the software. The technology also permits the user to share files. Instant messaging has become popular because the software implementing the service can be downloaded at no cost, and is easy to install and use.

1.50 Depending on the type of software used, the program will, when a message is initiated, connect the two devices, either via a direct point-to-point configuration or via a client-server configuration, through the ports of the devices. There are two

significant problems. First, in a client-server configuration, the instant message server may not necessarily log such messages, which means that such conversations can be considered conceptually similar to conversations over the telephone. Secondly, the program may have a feature that allows for messages to pass through legitimate open ports if others are not available. Whether such conversations are recorded will depend on the software used. In an earlier variation of Instant Messaging known as *Internet Relay Chat* (IRC),[1] conversations take place in a similar way to a conference call. IRC is mainly designed for group communications, though it also allows for one-on-one communications via private messages. It frequently suffers from the same issues as Instant Messaging, in that the servers relaying messages are not typically configured to log conversations.

1 'Internet Relay Chat' (*Wikipedia*) <https://en.wikipedia.org/wiki/Internet_Relay_Chat>; C Kalt, 'Internet Relay Chat: Client Protocol' (2000) RFC 2812 <https://tools.ietf.org/html/rfc2812>; and 'Internet Relay Chat: Server Protocol' (2000) RFC 2813 <https://tools.ietf.org/html/rfc2813>.

Peer to peer networking

1.51 As personal computers have developed, so have their capacity and power increased. As a result, there is less of a dividing line between a client and a server. This is because any host can be made a server by installing appropriate software into the computer. The software then permits other clients to obtain access to the resources of the computer over the network. This is called *peer-to-peer networking* (P2P),[1] and is often the subject of litigation regarding intellectual property, especially for the purpose of downloading music and films without payment. For instance, in Hong Kong, a *Norwich Pharmacal*[2] order was granted in the case of *Cinepoly Records Co Ltd v Hong Kong Broadband Network Ltd*[3] in respect of a number of IP addresses, and in the case of *Polydor Ltd v Brown*,[4] summary judgment was granted against the second defendant, Mr Bowles, for copyright infringement, after a *Norwich Pharmacal* order was made against various Internet service providers whose subscribers' IP addresses had been identified as being used for allegedly infringing activity. In both cases, the infringers were identified by the Internet service providers from their electronic records of the IP addresses assigned to their subscribers at the date and time in question when the allegedly infringing activity was taking place.[5]

1 Geoff Fellows, 'Peer-to-peer networking issues – an overview' (2004) 1 Digital Investigation 3; 'Peer-to-peer' (*Wikipedia*) <https://en.wikipedia.org/wiki/Peer-to-peer>; G. Camarillo (ed.), 'Peer-to-Peer (P2P) Architecture: Definition, Taxonomies, Examples, and Applicability' (2009) RFC 5694 <https://tools.ietf.org/html/rfc5694>.

2 *Norwich Pharmacal Co v Customs and Excise Comrs* [1974] AC 133 (CA), revd [1974] AC 133 (HL). See generally Paul Torremans, *Holyoak and Torremans Intellectual Property Law* (8th edn, Oxford University Press 2016), 612, 694–7.

3 [2006] HKCFI 84; [2006] 1 HKLRD 255; HCMP2487/2005 (26 January 2006).

4 [2005] EWHC 3191 (Ch).

5 For a similar case in Denmark, see Per Overbeck, 'The burden of proof in the matter of alleged illegal downloading of music in Denmark' (2010) 7 Digital Evidence and Electronic Signature Law Review 87; Per Overbeck, 'Alleged illegal downloading of music: the Danish Supreme Court provides a high bar for evidence and a new line of direction regarding claims for damages and remuneration' (2011) 8 Digital Evidence and Electronic Signature Law Review 165; similar comments were made by Baker DJ in *VPR Internationale v Does* 1-1017 2011 WL 8179128 (C.D.Ill. Apr. 29, 2011); Thomas M Dunlap and Nicholas A Kurtz, 'Electronic evidence in torrent copyright cases' (2011) 8 Digital Evidence and Electronic Signature Law Review 171.

Social networking

1.52 The advent of Web 2.0 has seen an enormous increase in websites that permit users to provide their own content. This varies in type from uploaded video clips (on sites such as YouTube), photographs (on sites such as Flickr), personal musings in the form of blogs (personal Web logs) and interactive exchanges with a wider audience in the form of social networking sites (such as Facebook and Twitter) and their more business-oriented alternatives (such as LinkedIn). As social networking has increased in popularity, with meteoric growth in participating users, several contexts arise in which the content of an individual's social network contribution may constitute evidence. For instance, an individual may be located at a specific place by means of his geotagged submissions to such a site, and photographs uploaded to a social networking site often retain their geotag data and reflect the time and place at which they were taken. Many of such sites with contributions that contain such information have been used for the purposes of grooming[1] and blackmail.[2]

1 *R v Lawrence Michael Scott* [2008] EWCA Crim 3201; *R v C.B.* [2010] EWCA Crim 3009.
2 *R v Jake Breakwell* [2009] EWCA Crim 2298.

1.53 In a different vein, an individual's social network contributions may suffice to determine political or social prejudices that in turn shed light on the character of a trial witness. The evidence in such cases may be recovered from the witness' contributions to the social networking sites, depending upon the availability and accessibility of such contributions to such sites. If an individual had made such contributions under an alias, a digital evidence professional may be able to establish his true identity by matching his online contributions to the same content that is found on the individual's storage media.

Concluding remarks

1.54 Given the ubiquity of digital devices and our near total reliance on them, the range of electronic evidence that is capable of being captured, investigated and disclosed in legal proceedings is very wide, as demonstrated in this chapter. From the files on a digital camera to the complex behaviour of a computer attached to the Internet, assessing electronic evidence has become the staple of a lawyer's life. Every lawyer should be equipped to offer appropriate advice to his clients in relation to the investigation, disclosure, admissibility and treatment of such electronic evidence. All these issues will receive due consideration in the subsequent chapters.

2

The characteristics of electronic evidence

Burkhard Schafer and Stephen Mason

2.1 Lawyers are required to offer appropriate advice to clients in relation to the disclosure or discovery of data in electronic form. If lawyers fail in their duty to more fully understand the issues surrounding digital data, they may find themselves subject to actions for negligence. Trying to persuade lawyers that they need to keep up to date with technology is far from new. In 1904, judges and lawyers were urged to make themselves aware of photography because 'they might otherwise accept what appears to be pure untouched work as reliable which was all the time outrageously worked on'.[1] And in 1959, an academic noted that 'hundreds of important cases involving disputed typewriting have been tried but there are still lawyers here and there who apparently have never heard of them and courthouses where a disputed typewriting has never been considered'.[2] Although written more than 50 years ago, the statement is undoubtedly still true today in many jurisdictions.

1 'Photographs as Evidence' (1903) 115 LT 474.
2 Winsor C Moore, 'The questioned typewritten document' (1959) 43 Minn L Rev 727, 727–8.

2.2 Electronic evidence and computer forensics are relatively recent additions to the means of proof in legal proceedings. Unlike many older forensic disciplines that were often introduced into the trial process with little or no legal debate and scrutiny, electronic evidence has caused considerable, and often controversial, discussion among legal professionals. Different legal systems have reacted in various ways to this new challenge. Some systems have introduced new legislation to specifically address electronic evidence. Other systems try to establish a 'closest match' to existing evidentiary concepts and have applied wherever possible existing rules analogously, for instance whether electronic evidence was admissible depended on whether it was similar to proof by (paper) document or proof by visual inspection. Most systems adopt a combination of both strategies. Where new legislation is introduced, the emphasis is on the differences between electronic and traditional forms of evidence. This can prevent lawyers from utilizing their collective institutional experience in evaluating and interpreting such evidence, often creating a sense of confusion and uncertainty. Where analogous approaches are used, the emphasis is on the similarities between traditional and digital evidence. Although this permits lawyers to draw on their experience in assessing the strength of the competing narratives that are argued by the parties, this can result in the inappropriate application of evidentiary rules. In either case, it is important for lawyers to be aware of the distinctive characteristics of electronic evidence to enable them to confidently and reliably evaluate the use of electronic evidence.

2.3 Defining what we mean by 'electronic' evidence is not an easy task. The type of evidence that we are dealing with has also been variously described as 'digital evidence' or 'computer evidence'. All three terms express some aspects of our pre-theoretical intuition that this type of evidence has some distinctive features that

Burkhard Schafer and Stephen Mason, 'The characteristics of electronic evidence', in Stephen Mason and Daniel Seng (eds.), *Electronic Evidence* (4th edn, University of London 2017) 18–35.

set it apart from other means of proof. However, defining what these distinguishing features are is far from straightforward. The rapid technological change in the field of information technology means that any definition narrowly tailored to the current state of technology faces the risk of becoming obsolete rapidly. Definitions that are suitably future proof by contrast tend to be too abstract and will cut across traditional divisions and categories in the law of evidence. For our purpose, we will take as our approach the need of the lawyer to turn certain artefacts – digital objects such as computer print-outs – into evidence that can be used for the purpose of proof in legal proceedings. Such a legal-purposive definition may not always map perfectly to the terminology in computer science, but if we keep this caveat in mind, we can develop a workable definition that will suit most applications and purposes.

2.4 Various definitions of electronic evidence exist. These include 'information of probative value that is stored or transmitted in binary form'[1] and 'information stored or transmitted in binary form that may be relied on in court'.[2] In his treatise, Casey defines digital evidence as:

> any data stored or transmitted using a computer that support or refute a theory of how an offense occurred or that address critical elements of the offense such as intent or alibi.[3]

1 Scientific Working Groups on Digital Evidence and Imaging Technology, 'Best practices for digital evidence laboratory programs glossary: version 2.7'.

2 International Organisation on Computer Evidence, *G8 proposed principles for the procedures relating to digital evidence* (IOCE 2000). This definition has been adopted by the US Department of Justice Office of Justice Programs, National Institute of Justice, in *Electronic Crime Scene Investigation: A Guide for First Responders* (US Department of Justice 2001) and *Forensic examination of digital evidence: A guide for law enforcement* (US Department of Justice 2004).

3 Eoghan Casey, *Digital Evidence and Computer Crime* (3rd edn, Academic Press 2011) 7.

2.5 Although the emphasis of this definition is on criminal investigations, it is a wider definition than the previous definitions, and it usefully explicates certain important aspects of electronic evidence. For instance, the reference to 'data' is to information that is held in electronic form, such as text, images, audio and video files. Also, the word 'computer' must be understood in its widest possible sense, and incorporates any device that stores, manipulates or transmits data. In addition, the definition implies that the evidence must be relevant and admissible, a question that can only be answered after we know what the electronic evidence, whether admissible or inadmissible, actually is.

2.6 With the aim of offering a wider-ranging definition that includes civil and criminal cases, we propose the following definition:

> Electronic evidence: data (comprising the output of analogue devices or data in digital form) that is manipulated, stored or communicated by any manufactured device, computer or computer system or transmitted over a communication system, that has the potential to make the factual account of either party more probable or less probable than it would be without the evidence.

2.7 This definition has three elements. First, the reference to 'data' includes all forms of evidence created, manipulated or stored in a device that can, in its widest meaning, be considered a computer.[1] This is used here in a non-technical sense meaning roughly 'a gathered body of facts'. Computer scientists often distinguish between 'data' and

'programs'. This distinction is not helpful for our purposes. In a copyright case, if a defendant has allegedly installed an unauthorized operating system, the presence of the system on his computer is electronic data for our purposes.[2] Second, the definition includes the various devices by which data can be stored or transmitted, including analogue devices that produce an output. Ideally, this definition will include any form of device, whether it is a computer as we presently understand the meaning of a computer, telephone systems, wireless telecommunications systems and networks, such as the Internet, and computer systems that are embedded into a device, such as mobile telephones, smart cards and navigation systems. Third, the definition restricts the data to information that is relevant to the process by which a dispute, whatever the nature of the disagreement, is decided by an adjudicator, whatever the form and level the adjudication takes. This part of the definition includes one aspect of admissibility – relevance only – but does not use 'admissibility' in itself as a defining criterion, because some evidence will be admissible but excluded by the adjudicator within the remit of his authority, or inadmissible for reasons that have nothing to do with the nature of the evidence. This could be, for instance, because of the way it was collected, such as violating privacy or professional privilege in the process that can result in rendering the evidence inadmissible. However, the definition of electronic evidence is limited to those items offered by the parties as part of the fact-finding process. This contextual, teleological aspect of the definition excludes, for instance, electronic documents that are created during a trial in a purely administrative capacity, such as email reminders of the date of the hearing sent to the parties by the court administrators. Of course, the very same data can become 'electronic evidence' if offered in an appeal to show that the information was not sent out in a timely fashion if this is part of the complaint.

1 Excluding though for the time being the human brain, which has also been compared to a computer.
2 Obviously, we also do not use 'data' in the way it is sometimes understood in telecommunications, where only digital, but not analogue information, is sometimes referred to as data.

2.8 A particularly important form of evidence in all developed legal systems is proof by document. Consequently, electronic documents are a particularly important form of electronic evidence.[1] They are also a particularly good example to illustrate some of the pertinent characteristics of electronic evidence. Because of the importance of documents for our daily life, and the way we handle them as folders, documents and photocopies, when dealing with electronic documents, many of the most important software applications intentionally mimic the 'look and feel' of traditional, paper-based stationery. We therefore create digital objects that are called documents, have the same visual appearance as documents typed on paper, 'turn' their 'pages' (as with some electronic readers for ebooks and ejournals), 'put' them in files and folders, and discard them in paper baskets. Email also intentionally mimics the traditional letter, from the letter icon on the inbox to the pencil icon to 'write' rather than type a new letter. This inauthentic familiarity can create the misleading impression that the electronic document exists somewhere on the computer as a single, complete whole and maintains its structural integrity even when the file is closed or the computer switched off, in the same way a paper document continues to exist when we put it out of sight and into a folder. This overly naive view underestimates the differences between electronic and paper based documents, and potentially also overestimates their reliability. The converse, however, can equally happen, where a more sophisticated user sees through the processes that intentionally create the appearance of a paper document, and

dismisses all electronic evidence as essentially deceptive, spurious, and unreliable rather than as a new kind of document. This becomes a particular problem for those jurisdictions whose evidence law has formal definitions of 'document' and proof by document, as for instance, the German 'Urkundenbeweis'. In these jurisdictions, legal rather than factual issues can increase the chasm between electronic and traditional documents and require bridging legislation necessary to make electronic documents also 'documents-in-law'.

1 William Kent, *Data and Reality* (2nd edn, 1stBooks 2000) for an interesting discussion of how humans perceive and process information, and how humans impose this outlook on data processing machines.

2.9 A better and more realistic approach is to acknowledge that documents in electronic form have particular characteristics that affect both the test for authenticity (or provenance) should authenticity be in issue, and the way the electronic evidence is secured and handled at the pre-trial stage. Arguably, evidence in electronic form ought to be subject to a more rigorous mechanism than would normally be associated with a document extant on physical media. John D. Gregory has observed that the integrity of physical documents is 'often protected fairly casually',[1] yet the same could not be said of documents that are created, modified, communicated, stored and deleted in electronic form. For instance, a forensic document examiner can analyse the chemical properties of the ink on a paper document to determine if more than one writing utensil was used, or if the ink is consistent with the time at which the document was allegedly created, or the material properties of the paper. Once the document is written, changes or alterations will also leave physical traces. With paper documents, we have therefore a clear understanding, routinely recognized in evidence law, between the original[2] and its copies. They are objects with different physical properties. This crucial distinction becomes problematic in the electronic medium, where not only copy and original are indistinguishable, but the very act of working on 'a' document will automatically and routinely without knowledge of the author create numerous copies on the computer, copies that can persist and record earlier drafts even when the document is completed. Documents in electronic form have a number of features that present particular challenges that a paper carrier does not in the physical world, as outlined below.

1 John D Gregory, 'Authentication rules and electronic records' (2002) 81 Can Bar Rev 529, 533.
2 For a short note on the meaning of 'original', see Stephen W Teppler, 'Digital data as hearsay' (2009) 6 Digital Evidence and Electronic Signature Law Review 7, 9 n 18; Stephen Mason, 'Electronic evidence and the meaning of "original"' (2009) 79 Amicus Curiae 26.

The dependency on machinery and software

2.10 Traditional documents make it easy for a reader to obtain access to information long after it was created with little or no additional costs. The only thing necessary is good eyesight, or a device to read the text to the person, and a knowledge of the language in which the document is written. This enables us to obtain access to information stored on ancient manuscripts and scrolls. Data in electronic form by contrast is dependent on hardware and software. The data requires an interpreter to enable it to be rendered into human-readable format. A user cannot create or manipulate electronic data without appropriate hardware. An electronic document should not be treated as an *object* 'somewhere there' on the computer, in the same way

as a paper book is in a library. Instead, the electronic document is better understood as a *process* by which otherwise unintelligible pieces of data that are distributed over the storage medium are assembled, processed and rendered legible for a human user. In this sense, the electronic document is nowhere: it does not exist independently from the process (software) that recreates it with the device (hardware) every time a user opens it on screen. If those electronic documents were produced in the 1990s, many thousands of these programs are now no longer available commercially, and even if such software were available, it might be impossible to load it on a modern operating system. An additional problem for older data is that it might be necessary to have a specific machine with specific software loaded in order to read the data.[1] This can cause additional expense to a party, as in the case of *PHE, Incorporated dba Adam & Eve v Department of Justice*,[2] where PHE was ordered to review information contained in a database, even though no program existed to enable it to obtain the information requested by the Department of Justice.

1 For instance, the jazz club Ronnie Scott's, based in Soho, London, was refurbished in 2005–6. As each part of the club was renovated, so large numbers of recordings of jazz musicians and singers, such as Dizzy Gillespie, Ella Fitzgerald, Chet Baker, Sarah Vaughan and Buddy Rich, recorded during live performances, were discovered. Some of the recordings were made on tapes that required machines that were no longer in the possession of the club. Report by Bob Sherwood, 'Ronnie Scott's jazz club to release archive of the greats', *Financial Times* (London, 28 June 2006) 1.

2 139 F.R.D. 249 (D.D.C. 1991); a similar problem was considered by Vinelott J in *Derby & Co Ltd v Weldon (No. 9)* [1991] 2 All ER 901, [1991] 1 WLR 652 (Ch).

The mediation of technology

2.11 Data in electronic form must be rendered into human-readable form through the mediation of a set of technologies. This means differences occur in how the same source object is displayed in different situations. A good example that is common to all users of the Internet is that a website can look very different depending on what type of screen and what browser is used, among other things. As a result, there can be no concept of a single, definitive representation of a particular source digital object. This can have obvious legal repercussions. An electronic contract document carelessly drafted may informally refer to the 'paragraphs' of the document without enumerating them since the formatting on the author's computer makes them plainly visible through line breaks in the text. Sent by email to the buyer and opened on her machine with a different software program, this formatting data may be unreadable and the paragraphs no longer apparent. Another example could be found in the changed representations of 'emojis' (ideograms used in an electronic message similar to older ASCII emoticons). For instance, in 2016, Apple controversially changed a 'hand gun' emoji into a 'water pistol' emoji. However, when a message containing this emoji is send to a non-Apple device, it could appear on the recipients' machine as a cartoon image of a real gun.[1] If a message such as 'bring <gun emoji> to our meeting' or 'retract that or I come with my <gun emoji>' is sent, what was intended by the sender as a light-hearted joke will look like a threat for some recipients, depending on what device they are using.

1 Bonnie Malkin, 'Water pistol emoji replaces revolver as Apple enters gun violence debate', *The Guardian* (London, 2 August 2016) <www.theguardian.com/technology/2016/aug/02/apple-replaces-gun-emoji-water-pistol-revolver-violence-debate>.

2.12 With traditional evidence, the act of observing or analysing a crime scene should not be allowed to alter it, a problem commonly known as 'contamination'. In contrast, with electronic evidence, the mere act of starting a computer and opening a document changes it, for instance, by altering its metadata. Different observers using only marginally different machinery will recreate different versions of the object in question, and it is not an easy issue to decide which one of them should be regarded as 'more authentic'.

2.13 To manage this issue, we can perhaps use the approach taken with eyewitness evidence. We know that different observers of the same event will always provide subtly different accounts as to what happened. Furthermore, an observer will unintentionally and inevitably alter his memory of the events every time he tries to remember them. In the same way in which we try to minimize these effects through appropriate protocols and procedures – for instance for a line-up of people that might include the accused or the interviewing of witnesses – protocols and procedures used by the digital evidence professional can minimize, but not eliminate, the distortion that the investigation creates. This means that it is crucial to identify appropriate standards, protocols, benchmarks and procedures and the relevant hardware and software, in relation to the management and use of any item of electronic evidence.

Speed of change

2.14 Technology changes rapidly in operating systems, application software and hardware. As a result, data in digital form may reach a point when they cannot be read, understood or used with new software or hardware. For instance, a software company may no longer produce software that is backward compatible or 'downward compatible' (where new versions of software are able to operate with older products). Technical obsolescence is a major problem that affects every aspect of the legal process, especially because the rate of change has now become so rapid.

2.15 The incessant speed of change has another consequence, again best explained by contrasting electronic evidence with traditional evidence. Eyewitnesses' identification evidence is one of the oldest, if not the oldest, form of evidence used in trial. Despite this, the way we elicit and interpret eyewitness evidence in legal proceedings has changed little over the centuries, and legal systems regularly keep culturally obsolete concepts such as the oath or dock identification for their ritual value. Fingerprint evidence is much younger, with little over a hundred years of forensic use. Since its inception, while the basics of the discipline have remained the same, important changes in the way in which we interpret fingerprint evidence have been made, as have the features that we look for when establishing a match. A fingerprint expert trained 90 years ago would probably need at least a refresher course. DNA evidence is younger still, but in its 30-year history, there have been considerable changes in the way in which DNA is collected, analysed and interpreted. An expert trained in the 1980s would require considerable retraining to be able to deal with current technology and equipment. For electronic evidence, the pace of change is faster still. This makes it all the more difficult to keep lawyers and other non-experts briefed of the relevant developments, and increases the reliance on experts. It also means that it is essential that an expert has up-to-date knowledge and receives constant training, which are more important

than 'experience' in this field. A related problem to the rapid change over time is the horizontal diversification of software and hardware. If a DNA expert analyses a blood sample, she need not know in advance the age, nationality or gender of the donor. By contrast, the digital evidence professional needs to know, and be trained for, the specific type of device and software that she is asked to analyse.

2.16 The ability of those investigating crimes, for instance, is also hampered by the speed at which the technology changes. In particular, obtaining relevant electronic tools to analyse a device forensically can be difficult for two reasons: first, the tools have yet to be devised, and second, because such tools can be expensive. In the case of *R v Hallam*,[1] Sam Hallam's conviction for three offences of murder, conspiracy to commit grievous bodily harm and violent disorder was quashed. One of the grounds of appeal was that Mr Hallam was in possession of two mobile telephones, one of which was a 3G telephone. Although the police seized both telephones, neither was the subject of forensic analysis. The defence did not seek to have them analysed either.[2] It was subsequently established that evidence stored on the 3G telephone that suggested that both Mr Hallam's alibi was probably correct, and that the memory of both Mr Hallam and his alibi witness were at fault as to the date they were together. The observations by Hallett LJ, delivering the judgment of the court, illustrate a naiveté in the prosecution's forensic investigation of the data. She said

> 65. … For reasons which escape us [the mobile phones] do not seem to have been interrogated by either the investigating officers or the defence team. We can understand why cell site evidence in relation to the use of the phones may have been of limited value given the close proximity of the masts, the various scenes, and the homes of those involved. However, given the attachment of young and old to their mobile phones, we cannot understand why someone from either the investigating team or the defence team did not think to examine the phones attributable to the appellant. An analysis of mobile phone evidence played a part in the investigation: see the schedule of calls between the co-accused to which we have already referred.
>
> …
>
> 67. One reason proffered for the failure to examine the phone was that in 2004 the Metropolitan Police did not have the technology in-house to examine 3G telephones. However, given our limited knowledge, we would have thought that even a cursory check might have produced some interesting results. Further, it might be thought that the appellant would have alerted his defence team to the fact that he had taken photographs on his new phone in the days before and after the murder which might have jogged his memory and helped establish his whereabouts.

1 [2012] EWCA Crim 1158.

2 This highlights the need for lawyers to ensure they are competent to practice, for which see in particular, Denise H Wong, 'Educating for the future: teaching evidence in the technological age', (2013) 10 Digital Evidence and Electronic Signature Law Review 16 and Deveral Capps, 'Fitting a quart into a pint pot: the legal curriculum and meeting the requirements of practice' (2013) 10 Digital Evidence and Electronic Signature Law Review 23.

2.17 Because the electronic evidence in the telephone supported the defendant's alibi and contradicted the eyewitnesses' testimony, which Hallett LJ had described as 'rock solid', the court concluded that this was a case of mistaken identity and acquitted the defendant.[1]

1 [2012] EWCA Crim 1158, [77].

Volume and replication

2.18 Electronic documents are easy to manipulate: they can be copied,[1] altered, updated, or deleted (and deleted in the electronic environment does not mean expunged). The integration of telecommunications and computers to form computer networks (such as wide area networks and the Internet) further allows data to be created and exchanged in far greater volumes than had hitherto been possible, and across physical and geographical boundaries. In essence, email, instant messaging and Internet communications are a duplicate and distributed technology.[2] Once computers are networked together in this fashion, an electronic document may be transmitted and numerous copies distributed around the world very rapidly. By way of example, in *AMP v Persons Unknown*,[3] the claimant's mobile telephone was stolen or lost. It was not protected with a password. A number of photographs were stored on the telephone, some of which were of an explicit sexual nature. Shortly after the telephone went missing or was stolen, digital images were uploaded on various social media websites, enabling others to download and share the images. Some of the social media sites removed the images when requested, but the images were seeded onto a Swedish BitTorrent node and continued to circulate. Ramsey J decided that the claimant was entitled to an interim injunction to prevent the distribution of the digital images, either by conventional downloading from a site or by downloading using the BitTorrent protocol. The injunction was granted in the following terms:

> 50. I therefore grant an interim injunction in the following terms against persons unknown being those people in possession or control of any part or parts of the files listed in Schedule C to the order who are served with this order:
>
> (1) shall immediately cease seeding any BitTorrent containing any part or parts of the files listed in Schedule C of this Order.
>
> (2) must not upload or transmit to any other person any part or parts of the files listed in Schedule C of this Order.
>
> (3) must not create any derivatives of any of the files listed in Schedule C of this Order.
>
> (4) must not disclose the name of Claimant (or any other information which might lead to her identification) or the names of any of the files listed in Schedule C of this Order.

1 Allegations of copying large numbers of electronic documents (around 56,000) formed part of the allegations in *Vestergaard Frandsen A/S v Bestnet Europe Limited* [2007] EWHC 2455 (Ch), which is a judgment in relation to an application by the defendants to strike out the action on the grounds that it was vexatious and an abuse of the process; George L Paul and Jason R Baron, 'Information inflation: can the legal system adapt?' (2007) 13 Rich J L & Tech 1.
2 Social media websites and sending text messages on mobile telephones and other devices were used to foment rioting in the UK in 2011: *R v Blackshaw and others* [2011] EWCA Crim 2312.
3 [2011] EWHC 3454 (TCC).

2.19 The ease of communication and replication of electronic documents has increased the potential volume of data that need to be identified to obtain relevant documents pertaining to litigation or the prosecution of a criminal offence. For instance, as part of the Enron investigation, the Federal Energy Regulatory Commission made public a dataset corpus containing 500MB of messages. Yet 'traditional' messages like these are a minuscule minority of all the electronic data (and potential evidence) that

is routinely created by machines, such as monitoring and routing Internet traffic. In addition to the sheer volume of this data, it poses the additional problem that in its raw form, it is not intelligible to humans – most of the data are instructions sent between and for the use by other machines. To turn them into evidence for legal proceedings requires a significant amount of translation, or 'sense making' by a suitably qualified expert.

2.20 To deal effectively with this amount of data, other computer tools such as data-mining software will routinely be required. These methods of analysis carry their own problems of accuracy, reliability, prejudicial effects and so on. Link analysis software, for instance, can create from this data a picture of a network that shows how people in the company formed communication circles that can be interpreted as the core of a conspiracy, simply as a result of the way in which the software arranges and visualises the information or other design choices not supported by the actual evidence.[1] On the other hand, other forensic disciplines routinely use scientifically validated sampling techniques.[2] At present, there is still a tendency not to use the same sampling protocols for at least some types of electronic evidence, in particular the type of data that can in principle be assessed directly by humans. This can force witnesses, such as police officers, to visually inspect potentially large amounts of disturbing illegal material. However, some jurisdictions have begun to use statistical methods of (electronic) evidence collection more systematically. 'Predictive coding' or 'technology assisted review' uses Bayesian probability theory and machine learning to scan electronic documents for data relevant to the case, and automatically identifies 'good candidates' for further examination by humans. Used mainly in civil electronic disclosure or discovery, it acquired approval from the courts in 2016.[3]

1 Cathleen McGrath, Jim Blythe and David Krackhardt, 'Seeing groups in graph layouts' (1996) 19 Connections 22.

2 If 300,000 suspicious pills are seized, only a small sample of them will be tested for being illegal drugs, and a statistical confidence value reported. Colin G G Aitken and David Lucy, 'Estimation of the quantity of a drug in a consignment from measurements on a sample' (2002) 47 J Forensic Sci 968.

3 *Pyrrho Investments Ltd v MWB Property Ltd* [2016] EWHC 256 (Ch); *Brown v BCA Trading Ltd* [2016] EWHC 1464 (Ch); Clive Freedman, 'Technology assisted review approved for use in English High Court litigation' (2016) 13 Digital Evidence and Electronic Signature Law Review 139.

2.21 The ability to transfer evidence rapidly can also create issues relating to jurisdiction. Many computer users now routinely upload all their files for back-up purposes to Internet-based providers. Business data may be processed using 'cloud computing' technology, which involves outsourcing the data to third party servers not owned and controlled by the company and possibly located all over the world, with each server holding at any time only pieces of the data.[1] On the other hand, the automatic uploading of data also means that the user of a device loses control over the information she has created. It can become increasingly difficult to delete, or rid oneself of information once it has been created on a device and the information is uploaded onto the 'cloud'.

1 Miranda Mowbray, 'The fog over the Grimpen Mire: cloud computing and the law' (2009) 6 Scripted Journal of Law, Technology and Society 133 <www.law.ed.ac.uk/ahrc/script-ed/vol6-1/mowbray.asp>.

Metadata

2.22 Metadata is, essentially, data about data. For instance, the metadata in relation to a piece of paper as a physical document may be:

> Explicit from perusing the paper itself, such as the title of the document, the date, the purported name of the person(s) who wrote it, who received it and the location of the document.
>
> Implicit, which includes such characteristics as the types of type (font) used, such as bold, underline or italic, the location of the document such as a coloured file to denote a particular type of document, and document labels that also act as pointers to allow the person using the document to deal with it in a particular manner, such as a confidential file, for instance.

2.23 All documents in electronic format will contain metadata in one form or another, including email communications, spreadsheets, websites and word processing documents. In fact, an electronic document has to have metadata to help interpret the purpose of the digital document. Such data can include, and be taken automatically from the originating application software, or supplied by the person who originally created the record. The list of information that is available includes, but is not limited to: when and how a document was created (purported time and date), the file type, the name of the purported author (although this will not necessarily be reliable[1]), the location from which the file was opened or where it was stored, when the file was last opened (purported time and date), when it was last modified, when the file was last saved, when it was last printed, the identity of the purported previous authors, the location of the file on each occasion it was stored, the details of who else may be able to obtain access to it, and, in the case of email, blind carbon copy (bcc) addresses.

1 For instance, where a document is revised on a number of occasions, on different computers and by different people, the name of the author will probably bear no resemblance to the authorship of the document. In *Crinion v IG Markets Ltd* [2013] EWCA Civ 587, the judgment of the trial judge, HH Judge Simon Brown QC, was taken word-for-word from the closing submissions of Mr Chirnside counsel for the claimant, written in a Word file. The trial judge adjusted the text, and the 'properties' file in the Word version of the judgment indicated that the 'author' was shown as 'SChirnside'. Also, the person originating a document may not use a new file, but begin the document by opening an old file, deleting the majority of the text, then creating the genesis of the new text; further, the name of the author may not be accurate if somebody other than the purported author logged on to a computer or system using the name of the person, and there may be occasions that a person uses software on their own computer that has been installed and registered in another name – although if the metadata is correct, it can directly lead to a killer that has murdered a number of people over a long period of time, as in the case of *The State of Kansas v Dennis L. Rader*, Case No. 05CR498, 2005, 18th Judicial District Court, Sedgwick County, Kansas. The defendant entered a plea of guilty before Waller J on 27 June 2005.

2.24 Because metadata is typically created automatically by the software and without knowledge of the user, it is therefore also more difficult to alter, manipulate or delete. Imagine that Alice writes a document on a computer. The software will add metadata that is associated with this document, for instance the time when the document was created. The file where this information is stored is the metadata that records the time of the event of writing. Since it is not an intentional creation by the author, but an automatic, software-generated artefact that is often invisible to the user, she may not know about this data, and even if she did, may not know how to alter or delete it.

2.25 However, it must be said that metadata is not infallible. Its interpretation requires the need to make assumptions about the environment in which they were created. If the time on the device was not accurate (for instance, a laptop flown across time zones without being adjusted for this, or the clock is slow, or has been deliberately changed), the recorded metadata will be false. Since the environment can in this sense 'lie', informed criminals can intentionally manipulate the data. For instance, experienced phishing attackers who use email will not only forge the sender's address of the emails they send, but manipulate the entire header to conceal the place from where the email originates. Finally, since metadata is the unintentional creation of information by the environment, examiners or other third parties who are operating in the same environment will also create metadata, and so potentially contaminate the evidence. A careless digital evidence professional, or an IT administrator of a company who was alerted to potentially illegal activity by an employee, can by the very act of opening and looking at the file create new metadata and overwrite the old (a new time when the document was, according to the computer, created), thereby erasing potentially useful metadata about the illegal activity such as the actual date and time it was committed.

Types of metadata

2.26 In broad terms, there are three main types of metadata:[1]

> (i) Descriptive metadata describe a resource for a particular purpose, such as a disclosure or discovery exercise. The metadata may include such information as title, key words, abstract and the name of the person purporting to be the author. To understand the history of the document more fully, it would be necessary to obtain information about how and when the system recorded the name of the purported author.
>
> (ii) Structural metadata describe how a number of objects are brought together. Some examples of structural metadata include 'file identification' (e.g. to identify an individual chapter that forms part of a book or report), 'file encoding' (to identify the codes that were used in relation to the file, including the data encoding standard used (ASCII, for instance), the method used to compress the file and the method of encryption, if used), 'file rendering' (to identify how the file was created, including such information as the software application, operating system and hardware dependencies), 'content structure' (to define the structure of the content of the record, such as a definition of the data set, the data dictionary, files setting out authority codes and such like) and 'source' (to identify the relevant circumstances that led to the capture of the data).
>
> (iii) Administrative metadata, which provide information to help with the management of a resource. Administrative data is further divided into rights management metadata and preservation or record-keeping metadata.

1 For more information on metadata, see *Dublin Core Metadata Initiative* <http://dublincore.org>; National Information Standards Organization, *Understanding Metadata* (NISO Press 2004) <www.niso.org/standards/resources/UnderstandingMetadata.pdf>; M. Day, *DCC Digital Curation Manual Instalment on Metadata* (UKOLN 2005) <www.dcc.ac.uk/resources/curation-reference-manual/completed-chapters/metadata>.

2.27 The metadata can be fundamentally linked to and be a part of the electronic document, included in the systems used to produce the document, or linked to it from a separate system.[1] Metadata can be viewed in a variety of ways, one of which is to

look at the 'properties' link in the application that created the document, or by using software specifically written for the purpose. Some metadata can also be removed with specialist software. This can be useful when sending files to third parties, but can attract additional expense if a court orders the data to be delivered up in its original format, as in the case of *Williams v Sprint/United Management Company*.[2] Before passing electronic spreadsheet documents in Excel form to the plaintiffs, Sprint modified the electronic files by, among others, deleting metadata from the electronic files that included the spreadsheets, and prevented the recipients from viewing certain data contained in the spreadsheets by locking the value of certain cells. Sprint was ordered to produce the spreadsheets in the manner in which they were maintained, including the metadata, although the adverse analyses and social security numbers could be redacted, and it was also ordered to produce unlocked versions of the spreadsheets. In his judgment, the judge discussed metadata and whether it formed a sufficient part of a document in electronic format for it to be given up to the other party.[3]

1 See also the discussion by Waxse J in *Williams v Sprint/United Management Company* 230 F.R.D. 640, 646–47 (D.Kan. 2005).
2 230 F.R.D. 640, 646–48 (D.Kan. 2005).
3 230 F.R.D. 640, 646–48 (D.Kan. 2005).

2.28 A further illustration of the importance of metadata is the case of *Campaign Against Arms Trade v BAE Systems PLC*.[1] Mr Justice King granted Norwich Pharmacal relief to the Campaign Against Arms Trade (CAAT) against BAE Systems PLC (BAE). On 29 December 2006, a senior officer of CAAT, Ms Feltham, sent an email to the members of the CAAT steering committee using an internal email list (caatcommiteee@lists.riseup.net), a private list not open to the members of the public and comprising only the 12 members of the steering committee and seven members of CAAT's staff. The email contained privileged legal advice that CAAT received from its solicitors. A copy of the email was somehow sent to BAE. By a letter dated 9 January 2007 and received the next day, solicitors for BAE returned a copy of the email printed on paper to CAAT's solicitors. This was the first time that CAAT came to know of the leak. The printed email returned to CAAT was incomplete (because the email metadata was missing). As described by Mr Justice King:

> It was a redacted version of that which had come into the possession of the Respondent and/or its own solicitors. All the routing information, the header address and so forth, which would give details of the email accounts through which the email had been received and sent before arriving at the Respondent and its solicitors, had been removed. Such removal must have been done either by the Respondent or by its solicitors acting on its instructons.[2]

1 [2007] EWHC 330 (QB).
2 [2007] EWHC 330 (QB), [31].

2.29 The source of the leak could only be the result of two possibilities, and CAAT did attempt, unsuccessfully, to trace the source, as described by Mr Justice King:

> 45. [T]here are really only two broad possibilities: either the source is one of the authorised recipients of the email, i.e. a member of the Applicant's steering committee or staff, or the email was intercepted or retrieved by other means by a person or persons unknown, be it by improper access to the Applicant's or a recipient's computer system, interception at [the email distribution list] or at some point whilst the email was sent over the Internet. In her first witness statement

> she explains how she made enquiries of each of the authorised recipients who each denied forwarding the email on. Her second witness statement was made in response to that part of the Respondent's skeleton argument in which it is said that the Applicant has not done enough and that before seeking the present order the Applicant should have … 'examined the electronic data available to it on its own computer systems and those of [the email distribution list] and further should have asked any authorised recipients to provide it with access to their personal electronic data for purpose of determining whether their denials of involvement in the copying are accurate'.
>
> 46. In this later statement Ms Feltham says she did check the 'sent folders' on the personal computers of the staff based in the Applicant's office, but explains that there was a major practical and logistical problem as regards access to the computers used by members of the steering committee. Unlike the staff they are not employees of the Applicant but volunteers who do not work in the office or use computer systems belonging to the Applicant. Some are members of other organisations who access emails from accounts and equipment owned by their employers. Some are based outside London. This all means that to have investigated further on the lines suggested by the Respondent, the Applicant would have needed access to computers to which the Applicant has no right of access and in any event the Applicant would have needed the 'costly services of a computer expert to go on a fishing expedition for emails which might or might not have been sent which moreover would have been very time consuming'.

2.30 The claim by BAE that CAAT ought physically to examine every computer to trace the route of the email is somewhat unrealistic, as explained above, and also fails to grasp the fundamental issue: that electronic data knows no geographical or physical bounds. Returning the email without the metadata is similar to returning a letter received through the post in an envelope, yet refusing to deliver up the envelope. That the routing and other technical data is 'similar' to the data included on an envelope is an understatement, because the routing and other metadata available in relation to an email is far more extensive than the metadata contained on an envelope. In this instance, Mr Justice King concluded that the order sought ought to be granted, although not in the terms requested.

2.31 This application illustrates the importance of the metadata associated with an electronic object. Documents in electronic form include metadata as a matter of course, and it seems unrealistic for the recipient to refuse to deliver up the full document, including the associated metadata, in such circumstances.

2.32 A case from the United States of America serves to highlight how concerns relating to the preservation of data are viewed, and the relevance of metadata. In the case of *Armstrong v Executive Office of the President, Office of Administration*,[1] researchers and non-profit organizations challenged the proposed destruction of federal records. The Executive Office of the President, the Office of Administration, the National Security Council, the White House Communications Agency, and the Acting Archivist of the United States intended to require all federal employees to print out their electronic communications on to paper to discharge their obligations under the provisions of the Federal Records Act. The members of the United States Court of Appeals, District of Columbia Circuit, rejected this solution, because in the words of Mikva CJ, the hard copy printed version 'may omit fundamental pieces of information

which are an integral part of the original electronic records, such as the identity of the sender and/or recipient and the time of receipt'.[2]

1 1 F.3d 1274 (D.C. Cir. 1993).
2 1 F.3d 1274, 1277 (D.C. Cir. 1993).

Social context and metadata

2.33 A significant amount of electronic data is created through communication between people separated by geographical, political, social and cultural boundaries. While the Internet brought people previously separated by distance into interaction, it also creates a new form of 'distance' between the communicators. Some communication practices do not translate well to this new medium, such as facial expressions and tone of voice. Evidence is not created in a vacuum, however. It has meaning, and can be interpreted only with knowledge of the context in which it was created. The exchange 'I hate you all and wish you were dead' between a teenager and his parents about cleaning a room will be interpreted by most people acquainted with a similar cultural background as insignificant and not serious. The same words found on a carefully written letter will carry a different meaning. Therefore, consideration has to be given to whether an email, a Twitter post, or an exchange on a discussion forum is more similar to a letter, or to a direct verbal excange.

2.34 Consider the case of *Chambers v Director of Public Prosecutions*.[1] Paul Chambers was a registered Twitter user with the handle '@PaulJChambers'. He was due to fly to Belfast from Doncaster Robin Hood Airport to meet another Twitter user, identified as '@Crazycolours', on 15 January 2010.[2] On 6 January 2010, Chambers became aware of problems at Doncaster Robin Hood Airport because of adverse weather conditions, and he and Crazycolours subsequently entered into the following exchange on Twitter:

> '@Crazycolours: I [Chambers] was thinking that if it does then I had decided to resort to terrorism'
>
> '@Crazycolours: That's the plan! I am sure the pilots will be expecting me to demand a more exotic location than NI'

1 *Chambers v Director of Public Prosecutions* [2012] EWHC 2157 (Admin).
2 The facts are taken from the judgment of Lord Judge LCJ in *Chambers v Director of Public Prosecutions* [2012] EWHC 2157 (Admin); Lilian Edwards, 'Section 127 of the Communications Act 2003: threat or menace?' (2012) 23 Computers & Law 21.

2.35 The court noted that in the context of the bad weather, these comments from Chambers seemed to be a reference to the possibility of the airport closing. No reply from Crazycolours was produced in court. Two hours later, when Chambers found out that the airport had closed, he posted the following message, available to the 600 or so followers of his Twitter postings:

> 'Crap! Robin Hood Airport is closed. You've got a week and a bit to get your shit together otherwise I am blowing the airport sky high!!'

2.36 On 11 January 2010, five days after the comments were posted, Mr Duffield, the duty manager responsible for security at Robin Hood Airport, found the comments as he was searching for tweets about the airport while off duty at home. He referred

the 'tweet' to his manager, Mr Armson, who regarded the comment as a 'non-credible' threat, partly because it featured Chambers' name, and because Chambers was due to fly from the airport in the near future. He passed this 'tweet' to the airport police, who took no action, but referred the matter on to the South Yorkshire police.

2.37 The South Yorkshire police arrested Chambers on 13 January while he was at work on suspicion of involvement in a bomb hoax, seven days after the offending message was 'tweeted'. Interviewed under caution, Chambers repeatedly asserted that this 'tweet' was a joke or meant to be a joke and not intended to be menacing. He said that he did not see any risk at all that it would be regarded as menacing, and that if he had, he would not have posted it. In interview he was asked whether some people might get a bit jumpy and responded 'yah. Hmm mmm'.

2.38 Chambers was charged with the offence of sending by a public electronic communication network a message of a 'menacing character' contrary to s 127(1)(a) and (3) of the Communications Act 2003 and found guilty. His appeal to the Crown Court in Doncaster was dismissed and on further appeal, the question was whether the words he used were a 'menacing message sent through a public communication medium' and thus in violation of s 127(1)(a) and (3) of the Communications Act 2003.

2.39 The ensuing prosecution showed just how difficult this determination can be. Some security officers at the airport were willing to dismiss it outright as 'venting', while others were concerned enough to inform the police. The court of first instance, applying an abstract, decontextualized dictionary definition of 'menace', convicted Chambers. On appeal, the members of the Court of Appeal noted, however, that '[b]efore concluding that a message is criminal on the basis that it represents a menace, its precise terms, and any inferences to be drawn from its precise terms, need to be examined in the context in and the means by which the message ws sent.'[1] The Court of Appeal reversed the decision of the lower court and allowed the appeal against conviction because it was posted as a conversation piece for Chambers's followers, drawing attention to himself and his predicament. It was not addressed to anyone at the airport or anyone responsible for public security. The communication was airing the grievance that the airport was closed when the writer wanted it to be open, and identified the person making the 'threat' in ample time for it to be reported and extinguished.

1 *Chambers v Director of Public Prosecutions* [2012] EWHC 2157 (Admin), [31].

2.40 For the Court of Appeal to consider the social context in which the electronic evidence was to be understood must be correct. The visual form in which this evidence appears may not be a true account of the social meaning that informed the users when the evidence was created. For instance, a tweet may look like a warning, but it is certainly not understood as such by the participants. Since judges and jurors will often have very different technological experiences, it is tempting to lead sociological or psychological evidence on these issues, but procedural rules on admissibility may well prevent this. These issues are, however, outside the expertise of the digital evidence professional, who is not in any position to offer any opinion about them.

Storage media

2.41 Generally, the media upon which electronic data are stored is fragile. Electronic storage media is inherently unstable, and unless the media is stored correctly, it can deteriorate quickly without showing external signs of deterioration. It is also at risk from accidental or deliberate damage and accidental or deliberate deletion.

2.42 Computers and systems now operate largely in a networked environment. The networked world comprises devices (MP3 players, computers, laptop computers, mobile telephones, personal digital assistants (PDAs), and tablets) linked by means of applications (facsimile transmissions, voice over Internet protocol (VoIP), email, peer to peer software, and instant messaging) that run over networks (the Internet, intranets, wireless networking, cellular networks, and dial up). The nature of this setup is that almost everything anybody does on a device that is connected to a network is capable of being distributed and duplicated with consummate ease. As a result, the same item of digital data can reside almost anywhere. The ramifications for lawyers and police officers are obvious. The relevant document may be available, but it might not be clear where it resides. This affects how a criminal investigation is conducted, and how much effort a party to a civil case will have to devote to find relevant documents for discovery or disclosure.

2.43 An example from the United States of America serves to illustrate some of the problems faced by a large organization in locating relevant documents in electronic format, especially historical email correspondence. Zubulake, a director and senior salesperson with UBS Warburg LLC, commenced legal proceedings for gender discrimination when she was dismissed from her job. Among others, she alleged that her manager Chapin treated her differently. She sought disclosure of UBS email communications to support her action.[1] The parties disagreed about the extent of the disclosure of emails, although it was not in dispute that email was an important means of communicating since each salesperson received approximately 200 emails each day. Securities and Exchange Commission Regulations required UBS to store emails. UBS used two storage methods: back-up tapes for disaster recovery and optical disks. This meant that there were three possible places that relevant email communications could be found: in files that were in use by employees, emails archived on optical disks, and emails sent to and from a registered trader (internal emails were not recorded) that were stored on optical storage devices. Ninety-four back-up tapes were identified as being relevant for the purposes of disclosure. UBS used a back-up program that took a snapshot of all emails that existed on a given server at the time the back-up was taken; namely, at the end of each day, on every Friday night and on the last business day of the month. Because emails were backed up intermittently, some emails were not stored, in particular where a user received or sent an email and deleted it on the same day. Scheindlin J determined that Zubulake was entitled to disclosure of the emails because they were relevant to her claim. UBS was ordered to produce all relevant emails that existed on the optical disks or its servers at its own expense, and from five back-up tapes selected by Zubulake. A consulting firm restored and searched the tapes for US$11,524.63. Additional expenses included the time it took lawyers to review the emails, which brought the total cost to US$19,003.43. Some 1,541 relevant emails were discovered. Fewer than 20 relevant emails were found on the optical disks. In July 2003, Zubulake made a further application for the remaining back-up tapes to

be restored and searched. UBS estimated that the cost would be US$273,649.39, and applied for the costs to be shifted to Zubulake. In considering the seven factor test (which is not relevant for the purposes of this particular discussion), the judge noted that a significant number of relevant emails existed on back-up tapes, and there was evidence that Chapin deleted relevant emails. Scheindlin J decided that Zubulake should pay 25 per cent of the cost of restoring the back-up tapes. UBS were required to pay all other costs.

1 *Zubulake v UBS Warburg LLC* 217 F.R.D. 309 (S.D.N.Y. 2003); *Zubulake v UBS Warburg LLC* 216 F.R.D. 280 (S.D.N.Y. 2003).

2.44 The purpose of describing this example is to illustrate the problems that multi-national organizations have in locating relevant evidence in electronic form. The nature of the distributed environment means that a range of practical problems have begun to emerge in determining what material needs to be disclosed or discovered to the other side. First, it is necessary to prevent the destruction of evidence, and then it is necessary to establish where the evidence is likely to be, before undertaking the exercise of sifting through the various sources to identify relevant documents. This will invariably require a party to locate where all back-up tapes are situated, whether held on the premises, with third parties in off-site remote storage or on individual computers, servers, in an archive or a disaster recovery system. The types of storage media that will need to be identified and located include tapes, disks, drives, USB sticks, iPads, laptops, PCs, PDAs, mobile telephones, pagers and audio systems (including voicemail), to name but a few.[1] The fragility and the ubiquity of electronic storage have made the modern day discovery exercise a formidable process.

1 Detective Inspector Simon Snell, Head of the High Tech Crime Unit in Devon and Cornwall, is reported to have indicated that criminals are using satellite navigation systems, games consoles and handheld computers to try and hide their activities; see 'Paedophiles using satnavs to store porn' (*TechRadar*, 23 January 2008) <www.techradar.com/news/computing-components/storage/paedophiles-using-satnavs-to-store-porn-207202>.

An intellectual framework for analysing electronic evidence

2.45 However, as we have seen, despite these differences, evidence in digital form shares important features with other types of evidence. Eyewitness evidence, forensic trace evidence such as DNA and proof by document can all provide the basis for analogical reasoning to determine the evidentiary value of an item of digital evidence, if we are aware of the limitations of this analogy. For instance, the human brain is more than a computer, yet at present only electronic, not eyewitness evidence is subject to expert testimony. The digital evidence professional, however, has a different job from that of a DNA analyst or a forensic entomologist and in particular he deals with mathematical abstractions rather the empirical objects. Therefore, his findings will not normally be in the form of matching probabilities or other quantifiable, generalised statements.[1] 'Universal' theories of evidence are regrettably either rare, or too abstract to be of much practical value. However, the 'hierarchy of propositions' promoted by the Forensic Science Service in the UK has the potential to provide such a framework which can also help to illuminate further the distinguishing features of electronic

evidence and what they mean for practice. We can only outline here what an extension of this scheme to electronic evidence could look like. We have already implicitly used some of their ideas, for instance, in the definition of electronic evidence. To interpret evidence, the digital evidence professional (or the judge) has to consider propositions that represent respectively the prosecution or defence, or the pursuer or defendant. Evidential weight can only be ascertained if the propositions from both sides are considered, and the increase or decrease in likelihood for both is considered. An illegal image of a minor on a computer, for instance, can only be evaluated if we know both the prosecution and defence's hypotheses. The defence might claim that the computer was bought second-hand and the image came from the previous owner. If this was the defence, then and only then would the metadata associated with the image that establishes when it was downloaded be crucial.

1 A potential problem for jurisdictions that follow the US decision in *Daubert v Merrell Dow Pharm.*, 509 U.S. 579 (1993) that requires that experts report confidence values and error rates, something that rarely applies in computer forensics.

2.46 The Forensic Science Service distinguishes three levels where these conflicting propositions can occur at different places in the analysis. Using the earlier example of the illegal image of a minor, on the level 1, we have the description of the offence, the possession of abusive images of a child. Here, the opposing propositions may be:

> A is in possession of an illegal image.
>
> A is not in possession of an illegal image.

On level 2, we find descriptions of activities:

> A downloaded the image.
>
> It is suggested that some earlier owner downloaded the image.

On level 3, we find propositions about sources. In our case, these would be:

> The image comes from the computer of A.
>
> It is suggested that the image comes from another source.

Ultimately, level 1 propositions propagate to level 3 propositions. The more intermediate steps, assumptions and inferences are necessary for this propagation process, the more remote a piece of evidence will be from the ultimate probandum on level 3. Several studies have shown, with examples, how this analysis can help in the evaluation of heterogeneous evidence, from eyewitnesses to DNA.[1] The nature of digital evidence, so our claim proposes, is that on a like-by-like comparison and allowing for the machine-mediated nature of electronic evidence, the evidence will be several steps further removed from the ultimate probandum when compared with traditional evidence. Questions on the origin of the illegal images, in particular, will have to be answered to determine, for instance, whether A downloaded the illegal image. An explicit inference is therefore needed to bridge the gap between the zeros and ones on a suspect's hard drive and the propositional claim that he was engaged in the activity of downloading those illegal images.

1 I W Evett, G Jackson and J Lambert, 'More on the hierarchy of propositions: exploring the distinction between explanations and propositions' (2000) 40 Science & Justice 3.

3

The foundations of evidence in electronic form

Stephen Mason and Daniel Seng

3.1 By taking into account the defining characteristics of the digital world, the use and admissibility of evidence in digital form have largely been accomplished through the definition and redefinition of legal concepts in the malleable rules of evidence. This chapter sets out to review the rules of evidence in the categorization, means of proof, treatment and weight given to electronic evidence, and an overview of the issues of hearsay, the treatment of software code as the witness, the presumption that computers are 'reliable', and authentication of electronic evidence, that will be covered in detail in the other chapters.

Direct and indirect evidence

3.2 'Judicial evidence is used to prove either facts in issue, or facts from which facts in issue may properly be inferred'.[1] Where evidence is used to prove the facts in issue, it is direct evidence. Where evidence is used to prove the facts from which facts in issue may be inferred, it is indirect evidence. If the facts in issue involve proving the existence of an electronic record, the electronic record itself constitutes direct evidence. Direct evidence refers to evidence which prove the facts in issue, and indirect evidence, or circumstantial evidence, is defined as evidence which prove facts which are relevant to the facts in issue. The existence of a physical object can be either direct evidence or indirect evidence.[2]

1 Colin Tapper, *Cross and Tapper on Evidence* (12th edn, Oxford University Press 2010) 20.
2 Tapper, *Cross and Tapper on Evidence* 30.

3.3 However, unless the existence, character or circumstance of the generation or storage of an electronic record is itself a fact in issue, it is more frequently the case that electronic evidence is used as indirect evidence to prove certain facts from which the facts in issue may be inferred. For instance, if an electronic record is adduced in evidence to show that A owes B a debt, the electronic record as indirect evidence only proves that there is a record that A owes B a debt, and it is necessary to make the additional inference that A actually owes B a debt.

3.4 That evidence takes electronic form has not been an impediment to its admissibility. Judges have admitted digital records of the product of mechanical devices and automatic recordings, photographs,[1] tape recordings,[2] automated film recordings of the movements of a ship as traced by radar,[3] microfilm,[4] print-outs of test results undertaken on a breath test machine,[5] video recordings[6] and computer print-outs.[7] The types and categories of electronic evidence are not closed.

1 *R v The United Kingdom Electronic Telegraph Company (Limited)* (1862) 3 F & F 73; 176 ER 33, where a photograph was admitted to show the nature of the surface of a highway in respect of an allegation of an obstruction; although photographs have to be verified on oath to be considered as

Stephen Mason and Daniel Seng, 'The foundations of evidence in electronic form', in Stephen Mason and Daniel Seng (eds.), *Electronic Evidence* (4th edn, University of London 2017) 36–69.

more than mere pictures; *Hindson v Ashby* [1896] 2 Ch 1 (CA) 21; *R v Tolson* (1864) 4 F & F 103; 176 ER 488, where a photograph was admitted in a case of alleged bigamy to illustrate oral testimony (Willes J commented in his summing up to the members of the jury: 'The photograph was admissible because it is only a visible representation of the image or impression made upon the minds of the witnesses by the sight of the person or the object it represents; and, therefore, is, in reality, only another species of the evidence which persons give of identity, when they speak merely from memory' – the jury subsequently entered a verdict of not guilty); D W Elliott, 'Mechanical aids to evidence' [1958] Crim LR 5; E. Goldstein, 'Photographic and videotape evidence in the criminal courts of England and Canada' [1987] Crim LR 384.

2 *Harry Parker v Mason* [1940] 2 KB 590; *R v Burr and Sullivan* [1956] Crim LR 442; *R v Ali (Maqsud)* [1966] 1 QB 688, [1965] 2 All ER 464, [1965] 3 WLR 229 (CA); for an example in Scotland, see *Hopes and Lavery v HM Advocate* [1960] Crim LR 566, 1960 JC 104, 1960 SLT 264.

3 *The Statute of Liberty Owners of Motorship Sapporo Maru v Owners of Steam Tanker Statute of Liberty* [1968] 2 All ER 195, [1968] 1 WLR 739 (PDAD).

4 *Barker v Wilson* [1980] 2 All ER 81, [1980] 1 WLR 884, (1980) 70 Cr App R 283 (DC), in respect of the Bankers' Books Evidence Act 1879.

5 *Castle v Cross* [1985] 1 All ER 87, [1984] 1 WLR 1372 (DC).

6 *Kajala v Noble* (1982) 75 Cr App R 149, [1982] Crim LR 433 (DC); *R v Grimer* [1982] Crim LR 674, 126 SJ 641 (CA); *R v Thomas (Steven)* [1986] Crim LR 682 (video recording of route taken made in lieu of maps and still photographs); *XXX v YYY and ZZZ* [2004] 1 RLR 137; *R v Nikolovski* (1996) 111 CCC (3d) 403.

7 *R v Wood (Stanley William)* (1983) 76 Cr App R 23, [1982] Crim LR 667 (CA) (the results of an automated analysis); *R v Sinha* [1995] Crim LR 68 (CA) (alteration of medical data recorded on a computer).

Evidence in both digital and analogue form

3.5 Although there are differences in form and format between the analogue or non-electronic version of an item of evidence and its electronic equivalent, if the differences are not material, courts will not reject electronic evidence in favour of other forms of evidence.

3.6 The differences may be material depending on the facts in issue: the alternative representations of data in digital form, in human readable form on a screen, or on a printed piece of paper, may become significant. In *Maynard*,[1] the trial magistrate declined to admit a print-out purporting to indicate the dates and times when the accused obtained access to data stored in the computer on the basis that not all of the data that were evident on the computer screen were fully replicated on the print-out. In a motion to review the magistrates' decision, Wright J upheld the magistrate's decision. The judge observed that if all of the data were relevant, the prosecution could have recorded the data on the screen by video.[2] In this case, it was demonstrated that the information recorded on the print-out was incomplete and not an accurate rendition of the data, and it did not just involve minor format changes, as the prosecution sought to contend.

1 (1993) 70 A Crim R 133; also cited as *Rook v Maynard* [1993] TASSC 137, (1993) 2 Tas R 97, (1993) 126 ALR 150.

2 [1993] TASSC 137. This was in 1993, before the introduction of computers into courts.

3.7 In contrast, in *New York v Rose*,[1] Morse J in City Court, City of Rochester, New York had to consider the use and admissibility of 'computer generated simplified traffic information tickets' or 'e-tickets'. The defendants moved for dismissal of the charges for driving while intoxicated because the State Police issued the charges in

the computer generated simplified information form rather than the multi-copy handwritten simplified traffic information form used across New York State. In a carefully reasoned judgment, Morse J set out how the system worked, and determined that the computer terminal used by the police generated each e-ticket with simplified traffic information for the defendants, printed duplicate originals of the e-ticket, and affixed the arresting officer's electronic signature to the e-ticket. Although there were minor format differences such as the colour and the number of sides on which the e-tickets were printed, these differences were not sufficient to persuade the judge that the e-tickets conformed substantially to a paper ticket. Thus, the motion for dismissal was denied.

1 11 Misc.3d 200 (2005), 805 N.Y.S.2d 506, 2005 N.Y. Slip Op. 25526.

3.8 A similar consideration arose in *Griffiths v DPP*,[1] where photographs taken with a speed camera on photographic film were admitted as evidence of a vehicle being driven at a speed greater than the speed limit. The evidence was also available in digital form, and the defence argued that the digital data should have been disclosed as well as the printed photographs. It was revealed that the camera technician had carried out a secondary check to confirm the speed of the vehicle on the digital files of the photographs. The judge indicated that the photographs were real evidence – they showed the times at which the vehicle was driven crossing a number of pre-measured lines painted on the road – and that using all this information it was perfectly possible to carry out the secondary check from the photographs themselves. It was not necessary to carry out the secondary check on the digital files. For this reason, it was held that whether the digital data was disclosed to the defendant was irrelevant.[2]

1 [2007] RTR 44.
2 [2007] RTR 44, [34].

Metadata and electronic evidence

3.9 However, there is a distinction between a document in digital form (and the content of the digital document as a print-out) and the metadata logically associated with the document in digital form. The metadata may be relevant, either as indirect evidence in relation to the document in digital form, or it may itself be relevant as direct evidence. For instance, when there are multiple versions of a digital document, the metadata as indirect evidence will enable the parties to identify the most relevant version of the document. On the other hand, where there is an allegation that the user manipulated the metadata of the file such as its date-time stamp to his own advantage, the correct date and time of the file becomes the fact in issue and the metadata is the direct evidence. In such a case the metadata may need to be rendered into human-readable form.

Means of proof

3.10 All direct and indirect evidence used to prove a fact in issue or a relevant fact takes one (or more) of the following forms: testimony, hearsay, documents and real evidence.

Testimony and hearsay

3.11 Testimony is the declaration (which must be admissible) in court of a person who actually perceived the fact in issue or facts from which facts in issue may properly be inferred.[1] Thus the human perception of a computer display as narrated via oral testimony is admissible as evidence that a counterfeit computer game was being played in breach of copyright.[2]

1 The only exception to this general rule is the evidence of experts testifying to matters calling for their expertise. See Tapper, *Cross and Tapper on Evidence* 54.
2 The image on a screen can constitute sufficient evidence of data copied on to the RAM of a computer used to play counterfeit games to establish an offence of breach of copyright, for which see *Gilham v The Queen* [2009] EWCA Crim 2293.

3.12 If, however, the best that a witness can do is to depose as to what someone else said on the fact in issue, it will be hearsay, because it is 'an assertion other than one made by a person while giving oral evidence in the proceedings … as evidence of any fact asserted'.[1] In the context of digital evidence, what someone else said is typically recorded electronically. Hearsay is generally inadmissible unless it falls within one of the exceptions to the rule against hearsay. (A further treatment of this subject is found in the chapter on Hearsay.)

1 *R v Sharp* [1988] 1 All ER 65, 68, [1988] 1 WLR 7, 11.

Real evidence

3.13 The term 'real evidence' tends not to be used in practice,[1] and is best described as 'Material objects other than documents, produced for inspection of the court'.[2] Professor Smith considered that there is no authoritative definition of 'real evidence', and suggested that 'where a document is tendered simply to prove the fact that a statement was made (and not to prove a fact stated therein), it is not properly described as "real evidence"'.[3] Cross and Tapper, on the other hand, suggested that there is 'general agreement' that 'real evidence' covers the production of material objects for inspection by the judge or jury in court to reach its own conclusions on the basis of its own perception.[4]

1 Tapper, *Cross and Tapper on Evidence* 49. It was used in *O'Shea v City of Coventry Magistrates' Court* [2004] EWHC 905 (Admin).
2 Hodge M Malek (ed.), *Phipson on Evidence* (18th edn, Sweet & Maxwell 2013), paras 1–14.
3 *R v Spiby*, [1991] Crim LR 199 (CA) 202.
4 Tapper, *Cross and Tapper on Evidence* 49.

To highlight the difference between real evidence and hearsay in electronic evidence, Professor Daniel Seng and Sriram S. Chakravarthi formulated the following categorization: digital data that is stored on a device; a device that processes data, and a device that processes and stores data.[1] The first is hearsay, because the device is a record of human assertions. As for the second and third devices, where the data is produced without human intervention, it is real evidence. If the data is a record of human assertions, it is hearsay. Although the distinction is a clear one, it can be difficult to apply in practice,[2] as the following cases illustrate.

1 Daniel Seng and Sriram S Chakravarthi, *Computer Output as Evidence – Consultation Paper* (Singapore Academy of Law 2003) 87–8, available at <www.agc.gov.sg/DATA/0/Docs/PublicationFiles/Sep_03_ComputerOutput.pdf>.
2 Seng and Chakravarthi, *Computer Output as Evidence* 137–8; a point made by Adam Wolfson, '"Electronic fingerprints": doing away with the conception of computer-generated records as hearsay' (2005) 104 Mich Law Rev 165.

Evidence in analogue form

3.14 The treatment of evidence in analogue form (which preceded the use and acceptance of digital computers) first received detailed treatment in the case of *R v Ali (Maqsud)*[1] where the issue was the admissibility of a tape recording. In admitting the evidence, Marshall J analogized tape recordings with photographs, and noted that just as evidence of things seen through telescopes or binoculars which otherwise could not be picked up by the naked eye have been admitted, the same would apply to devices for picking up, transmitting, and recording conversations, but noted:

> [I]t does appear to this court wrong to deny to the law of evidence advantages to be gained by new techniques and new devices, provided the accuracy of the recording can be proved and the voices recorded properly identified; provided also that the evidence is relevant and otherwise admissible, we are satisfied that a tape recording is admissible in evidence.

1 [1968] 2 All ER 195.

3.15 Shortly thereafter, Sir Jocelyn Simon P determined in *The Statute of Liberty, Sapporo Maru M/S (Owners) v Steam Tanker Statute of Liberty (Owners)*,[1] that the film recording of a radar set of echoes of ships within its range was real evidence, even though it was recorded from a mechanical instrument.[2] The judge considered that there was no distinction in the manual operation of a camera by a photographer or the observations of a barometer operator and its equivalent operation by a trip, a clock or a dial recording mechanism. It held that 'the law is bound these days to take cognisance of the fact that mechanical means replace human effort',[3] and accepted that the film comprised real evidence because it recorded the information given out by the radar set, rejecting the submission that the evidence was hearsay.

1 [1968] 2 All ER 195.
2 Oral evidence of the position of a ship as given by a radar is acceptable, for which see *Chen Yin Ten v Little* (1976) 11 ALR 353.
3 [1968] 2 All ER 195, 196.

Evidence in digital form

3.16 The characterization of evidence as real evidence or as hearsay becomes more complicated with evidence in digital form, especially when some computational processing is made. In *R v Pettigrew*[1] the Court of Appeal held that the print-out from a computer operated by an employee of the Bank of England was a hearsay statement. The operator fed bundles of bank notes with consecutive serial numbers into the machine, and the machine automatically rejected any notes in the bundle that were defective. The machine also recorded the first and last serial numbers of each bundle of 100 notes. (As the operator fed the bundles into the machine, he also noted the first serial numbers in the bundle on a card.) It is the print-out from this machine that

was sought to be admitted in evidence. The purpose of adducing the evidence was to permit the prosecution to trace the issuance of the notes, and to link bank notes found in the possession of Pettigrew to a particular bundle of notes that had been stolen in a burglary. Counsel for the prosecution argued that the print-out was admissible under the provisions of the Criminal Evidence Act 1965 as a business record.[2] However, s 1(1)(a) required that for such a record to be admissible as evidence of the truth of any matter dealt with in the record, the information would have to be supplied by a person who had, or may reasonably be supposed to have, personal knowledge of the matters. The members of the Court of Appeal reached the conclusion that the operator did not have personal knowledge of the numbers of the notes that were rejected, because the machine automatically compiled the list.

1 (1980) 71 Cr App R 39; applied in *R v Wiles* [1982] Crim LR 669.
2 The Criminal Evidence Act 1965 was repealed by the Police and Criminal Evidence Act 1984 sch 7 pt III.

3.17 While this was an accurate application of the hearsay rule, the analysis omitted any consideration that the print-out might be considered real evidence.[1] Professor Smith noted that 'the operator had personal knowledge of the first number of each bundle which he fed into the machine because he recorded that number on a card',[2] and suggested that because the operator had knowledge of the number at a given point in time, it was not material that he forgot it. Once the first number could be established, it could then be inferred that the new notes bore consecutive serial numbers.[3] Professor Smith considered that this is not hearsay but direct evidence, because there was an absence of human intervention.[4] On the other hand, Professor Tapper took the view that the print-out was partly hearsay and partly non-hearsay – the first number is the hearsay and the last number and the numbers of the notes that were rejected were not hearsay because it was the output of the device.[5]

1 Colin Tapper, *Computer Law* (4th edn, Longman 1989) 375; print-outs were admitted under the provisions of s 1(1) of the Criminal Evidence Act 1965 in *R v Ewing* [1983] QB 1039, [1983] 2 All ER 645, [1983] 3 WLR 1 (CA), although Seng and Chakravarthi (n 1, 3.14) 90, point out that 'the electronic records *are* the manifestation of the transaction'.
2 J C Smith, 'The admissibility of statements by computer' [1981] Crim LR 387, 388.
3 *R v Pettigrew* (1980) 71 Cr App R 39, 42. In effect, Professor Smith's point was an argument pursued by counsel for the Crown.
4 Smith (n 2) 387 [389–90].
5 Colin Tapper, 'Reform of the law of evidence in relation to the output from computers' (1995) 3 Intl J L & Info Tech 87.

3.18 Professor Seng considered that the views of Professors Smith and Tapper were both plausible: 'The difference lies in whether the operator fed the first number into the machine, and whether the machine processed this number.'[1] Seng continued:

> … the different views espoused by Professors Tapper and Smith can be resolved as follows: was the machine operating as a data storage device in relation to the first number, or a data processing device? Some form of hybrid function may also be possible, eg, the operator inputs the first number, which the machine records and then verifies against its own reading of the first number. If the machine behaved in this way, perhaps Professor Smith's view is perhaps more accurate. This is all a question of the degree and extent of human intervention.[2]

1 Daniel K B Seng, 'Computer output as evidence' [1997] Sing JLS 139.
2 Seng, 'Computer output as evidence' 140.

3.19 As computers are designed to undertake a wide range of tasks, this means that the evidence available as an output of a computer is equally as varied. A review of the cases shows that whether electronic evidence is real evidence or hearsay turns on characterizing the evidence as being due either to a device's processing functions or to its storage functions.

3.20 In *Wood (Stanley William)*,[1] the computer was considered as a tool, and the print-out was an item of real evidence. The basis of admitting a print-out of an output as an item of real evidence was explained by Professor Tapper:

> Evidence derived from a computer constitutes real evidence when it is used circumstantially rather than testimonially, that is to say that the fact that it takes one form rather than another is what makes it relevant, rather than the truth of some assertion which it contains.[2]

1 (1982) 76 Cr App R 23. See also the earlier case of *R v McCarthy (Colin Paul), R v Warren (Mark Stephen), R v Lloyd (Leigh Cedric), R v Warren (Robert John)* [1998] RTR 374 (CA).
2 Tapper, 'Reform of the law of evidence in relation to the output from computers' 373.

3.21 The same distinction was drawn by Professor Smith as regards the computer print-out in *R v Ewing*[1] between its use as evidence to prove that a thing was done (money had been credited to a bank account), and evidence that something was recorded as being done (the bank clerk records a payment, as opposed to creating the credit).[2]

1 *R v Ewing* [1983] QB 1039, [1983] 2 All ER 645, [1983] 3 WLR 1 (CA).
2 [1983] Crim LR 472 (CA), 473.

3.22 The admissibility of more complex electronic evidence is illustrated in the case about the breath alcohol print-out from a portable measuring device, the Intoximeter 3000. In *Castle v Cross*,[1] it was determined that the print-out is an item of real evidence and not hearsay.[2] The judge compared the device to a speedometer, a calculator, or a sophisticated tool. In this instance, the breath alcohol value in the print-out comprised information that was produced by the Intoximeter, because the data had not passed through a human mind. On the other hand, Kennedy J also remarked that 'where a computer is used in respect of its memory function, it is possible to envisage where it might fall foul of the rule against hearsay.'[3]

1 [1984] 1 WLR 1372 (DC), 1380.
2 The members of the Court of Appeal in Northern Ireland followed this line, admitting a copy of a print-out as being real evidence in *Public Prosecution Service v Duddy* [2008] NCIA 18, [2009] NI 19.
3 [1984] 1 WLR 1372 (DC), 1380.

3.23 In *R v Spiby (John Eric)*,[1] Taylor LJ held that there was a distinction between a print-out as real evidence and as hearsay. Professor Smith[2] noted the difference between the content of the print-out as a mere recording of a fact, such as when data are processed by a computer without any human input of any description,[3] and the content of the print-out as being processed in some way by a human being. The print-out was generated by a computerized machine called a 'Norex', which monitored the telephone calls of hotel guests in order to work out how much to charge for the use of the telephone. It was held to be real evidence.

1 (1990) 91 Cr App R 186, 192, [1991] Crim LR 199 (CA).

2 Smith, 'The admissibility of statements by computer' 387.
3 Although no computer works on this basis – the code is written in the main by human beings, and the code comprises the instructions to the computer, upon which basis the computer undertakes activities, and the computer undertakes actions based on the instructions written by human beings.

3.24 In *R v Robson, Mitchell and Richards*,[1] a print-out of telephone calls made on a mobile telephone was adduced as evidence of the calls made and received in association with the number. The defence's challenge that the evidence was documentary hearsay failed. Orde J held that 'where a machine observes a fact and records it, that record states a fact. It is evidence of what the machine recorded and this was printed out … The record was not the fact, but evidence of the fact.'[2]

1 [1991] Crim LR 362.
2 [1991] Crim LR 362, 363; see also *McDonald v R* [2011] EWCA Crim 2933 where a print-out of telephone calls was admitted in the absence of the electronic records that no longer existed. Records of calls made by a mobile telephone were accepted as real evidence by the Court of Criminal Appeal of the Republic of Ireland in *People v Colm Murphy* [2005] 2 IR 125 (CCA) and in *People v Brian Meehan* [2006] 3 IR 468 (CCA).

3.25 In the business context, two popular uses of computers are the formation of records, and the recording of the credits and debits of an account. Where it is the latter, the records of computer payment transactions are considered real evidence, as their Lordships made clear in *R v Governor of Brixton Prison, ex p Levin*.[1] In this appeal against extradition, it was alleged that Levin used a computer terminal in St Petersburg to gain unauthorized access to a Citibank terminal in Parsipanny, New Jersey to make 40 fraudulent transfers of funds from the accounts of clients of the bank to accounts which he or his associates controlled. Print-outs of screen displays of the historical records of computer payment transactions were adduced, and a witness gave evidence as to how the records were created. Lord Hoffmann took the opportunity to make clear the difference between a hearsay statement and evidence of a record of a transaction:

> The print-outs are tendered to prove that such transfers took place. They record the transfers themselves, created by the interaction between whoever purported to request the transfers and the computer program in Parsipanny. The evidential status of the print-outs is no different from that of a photocopy of a forged cheque.[2]

1 [1997] AC 741, [1997] 3 All ER 289, [1997] 3 WLR 117 (HL).
2 [1997] AC 741 (HL), 746.

Documents and disclosure or discovery

3.26 'A document may be put in evidence either as a chattel … or else as a statement.'[1] If it is a chattel, it is admissible as real evidence as 'a substance such as a paper or parchment bearing an inscription'.[2] If it is a statement, it is admissible as testimonial evidence.[3] In such a case, the hearsay rules may apply to exclude the statement from admissibility, unless it falls within a hearsay exception.

1 Tapper, *Cross and Tapper on Evidence* 55–6.
2 Tapper, *Cross and Tapper on Evidence* 55–6.
3 Tapper, *Cross and Tapper on Evidence* 55–6.

3.27 It is in both contexts that in evidentiary discovery (or disclosure as it is now called in England & Wales), a 'document' has been construed widely. While the emphasis is on the recording of the content by the application of (usually text) on to (usually) paper, early decisions such as the Court of Appeal in *Lyell v Kennedy (No. 3)*[1] have admitted photographs of tombstones and houses as documents for the purposes of discovery. In *R v Daye (Arthur John)*,[2] Darling J suggested that the meaning of a document should not be defined in a narrow way:

> But I should myself say that any written thing capable of being evidence is properly described as a document and that it is immaterial on what the writing may be inscribed. It might be inscribed not on paper, but on parchment; and long before that it was on stone, marble, or clay, and it might be, and often was, on metal. So I should desire to guard myself against being supposed to assent to the argument that a thing is not a document unless it be a paper writing. I should say it is a document no matter upon what material it be, provided it is writing or printing and capable of being evidence.[3]

1 (1884) 50 LT 730; for a discussion about the status of legal resources on the Internet, included case reports, see R J Matthews, 'When is case law on the web the "official" published source? Criteria, quandaries, and implications for the US and the UK' (2007) 2 Amicus Curiae 19, 25.
2 [1908] 2 KB 333 (KBD).
3 [1908] 2 K.B. 333 (KBD), 340; see Malek (n 2, 3.13) para 41-02 for a more detailed discussion of documents within the rule.

3.28 In *Hill v R*, Humphreys J held 'that a document must be something which teaches you something … To constitute a document, the form which it takes seems to me to be immaterial; it may be anything on which the information is written or inscribed – paper, parchment, stone or metal'.[1] Likewise, statutes adopt a similarly broad definition of a 'document'. Section 13 of the Civil Evidence Act 1995 defines a 'document' as 'anything in which information of any description is recorded'. The same definition is provided in s 20D(3) of the Taxes Management Act 1970.

1 [1945] 3 KB 329, 332–3.

3.29 Audio tapes were accepted by Walton J as a discoverable document in *Grant v Southwestern and Country Properties Ltd*,[1] where a 'document' was defined as its quality to convey information. Television film is also considered a document,[2] as is the output of facsimile transmissions,[3] and a label on a bottle containing a specimen of blood provided by the accused.[4]

1 [1975] Ch 185, [1974] 2 All ER 465, [1974] 3 WLR 221. See also *R v Senat, R v Sin* (1968) 52 Cr App R 282; *R v Stevenson* [1971] 1 All ER 678, [1971] 1 WLR 1; *R v Robson (Bernard Jack); R v Harris (Gordon Federick)* [1972] 2 All ER 699, [1972] 1 WLR 651 (CCC).
2 *Senior v Holdsworth Ex p Independent Television News* [1976] QB 23, [1975] 2 All ER 1009, [1975] 2 WLR 987 (CA).
3 *Hastie and Jenkerson v McMahon* [1991] 1 All ER 255, [1990] 1 WLR 1575, (CA).
4 *Khatibi v DPP* [2004] EWHC 83 (Admin).

3.30 In *Derby v Weldon (No. 9)*,[1] one of the earliest modern decisions on the point, it was held that data stored on a computer in the form of an online database constitutes a document for the purposes of the obligation to discover under the provisions of Order 24 of the Rules of the Supreme Court. In analysing this point, Vinelott J referred to the Australian case of *Beneficial Finance Corp Co Ltd v Conway*,[2] in which McInerney J held

that a tape recording was not a document because the information is not capable of being visually inspected. Vinelott J however preferred the opposing view in *Grant v Southwestern and County Properties Ltd*,[3] in which Walton J pointed out that there is no difference between recording a conversation on a tape recorder and in shorthand. Both are methods of recording the same conversation. Vinelott J quoted Walton J with approval as follows:

> … the mere interposition of necessity of an instrument for deciphering the information cannot make any difference in principle. A litigant who keeps all his documents in microdot form could not avoid discovery because in order to read the information extremely powerful microscopes or other sophisticated instruments would be required. Nor again, if he kept them by means of microfilm which could [not] be read without the aid of a projector.[4]

1 [1991] 2 All ER 901, [1991] 1 WLR 652 (CA).
2 [1970] VR 321.
3 [1975] 1 Ch 185, [1974] 3 WLR 221, [1974] 2 All ER 465, 118 SJ 548 Ch D; Walton J criticised the reasoning of McInerney J at 196F–197A.
4 [1991] 2 All ER 901 (CA), 906B-C.

3.31 Thus the interposition of a computer to enable the retrieval of data stored in the online database did not disqualify the data from being considered a document. A similar issue as to the meaning of a 'document' in the context of data stored on a computer for discovery was also discussed in *Alliance & Leicester Building Society v Ghahremani* on a motion to commit Naresh Chopra, a solicitor, to prison for contempt of court.[1] Mr Chopra was alleged to have deliberately deleted part of a file that showed crucial transaction details stored on his computer in contempt of court, when investigations into possible mortgage fraud and negligence were being conducted into his affairs. A court order had directed Chopra to restrain from destroying or altering any document relating to the transaction, and required him to deliver up all such documents in his control. In the contempt proceedings, counsel argued that the word 'document' required there to be some form of visible writing on paper or other material, and because there was no physical document, the order had not been breached. Hoffmann J noted the comments of Vinelott J in *Derby v Weldon (No. 9)*,[2] and held that 'document' would bear the same meaning in the discovery order. Taking into account the expert evidence, Hoffmann J concluded that it was proved beyond reasonable doubt that Chopra did alter or destroy part of the file as a document,[3] and granted the motion, although Chopra was eventually fined instead.[4]

1 (1992) 32 RVR 198, [1992] TLR 129 (Ch).
2 [1991] 2 All ER 901, [1991] 1 WLR 652 (CA).
3 (1992) 32 RVR 198, 203. Forged evidence has increased. For some examples in the context of England & Wales, see *ISTIL Group Inc v Zahoor* [2003] EWHC 165 (Ch), [2003] All ER 252 [106]-[111] for a forged document; *Fiona Trust & Holding Corporation v Privalov* [2010] EWHC 3199 (Comm) [1405]–[1430] for a forged and back-dated agreement and employment contract; *Apex Global Management Ltd v FI Call Ltd* [2015] EWHC 3269 (Ch) for forged emails; in the criminal context, see *R v Brooker* [2014] EWCA Crim 1998 (available in the LexisNexis electronic database), where Brooker sent text messages from a second mobile telephone in her possession, claiming that her boyfriend sent them; *Islamic Investment Company of the Gulf (Bahamas) Ltd v Symphony Gems NV* [2014] EWHC 3777 (Comm) a case of fictitious litigation; *Otkritie International Investment Management Ltd v Urumov (Rev 1 - amended charts)* [2014] EWHC 191 (Comm), in which the allegations (and counter-allegations) included, amongst other things, the forgery of the contents of a laptop and metadata in relation to documents; Steven Morris, 'Barrister becomes first to be jailed for perverting justice', *The Guardian* (London, 20 September 2007).

4 Communications by email between Nicholas Leviseur, counsel for Mr Chopra, and Stephen Mason dated 14 October 2006 and 23 November 2006.

3.32 There is judicial recognition that the acceptance and use of technology will increase the range of objects that fall within the definition of 'document'. In *R v McMullen*,[1] Linden J held that a current account ledger card printed from a computer was a document within the meaning of s 29(2) of the Canada Evidence Act. The judge commented that: 'It is merely a new type of copy made from a new type of record. Though the technology changes, the underlying principles are the same.'[2] Citing this comment, Morden JA observed that the 'section should be considered as "always speaking" and "be applied to the circumstances as they arise …"'.[3] The same view was emphasized by Buxton LJ in *Victor Chandler International v Customs and Excise Commissioners*,[4] where he observed that '… the word "document" is not constrained by the physical nature that documents took in 1952, so we are entitled, and indeed bound, to consider the appropriate application of the concept of circulation, etc, of a document in the light of current practice and technology'. In this case, an advertisement contained in a teletext transmission was held to be a document for the purposes of the Betting and Gaming Act 1981. This view was reinforced by Pumfrey J in *Marlton v Tectronix UK Holdings*,[5] when the judge held that a computer database, in as far as it forms part of the business records of a company, is a document for the purposes of the Civil Procedure Rules, and therefore can be disclosed. Calvert Smith J also concluded, in *Kennedy v Information Commissioner*,[6] that the word 'document' in s 32 of the Freedom of Information Act 2000 included information recorded in an electronic medium. The judge said:

> It seems clear to me that for the Act to work at all – and in particular for Section 32 to work at all – the word 'document' must now mean what everybody now thinks it means and includes both hard and electronic copies of documents.[7]

1 42 CCC (2d) 67.
2 42 CCC (2d) 67.
3 *R v McMullen* (1979) 100 DLR (3d) 671, 676.
4 [2000] 2 All ER 315, 329.
5 [2003] EWHC 383 (Ch), [2003] Info Tech LR 258, 2003 WL 1610255.
6 [2010] EWHC 475 (Admin), [2010] 1 WLR 1489.
7 [2010] EWHC 475 (Admin), [79].

3.33 As such, a 'document' is a medium upon which information is stored. The medium may sometimes determine the admissibility of the evidence, but the definition of a document is considered wide enough to bring any medium into its ambit without causing difficulties.[1] This must be correct, because if information is not stored on a medium, the content is not available without the medium, and therefore the information remains oral evidence. As Lord Milligan in *Rollo (William) v HM Advocate*[2] said, when he indicated that the information stored in a Sharp Memomaster 500 hand-held device was a document:

> Unsurprisingly, the word 'document' in normal usage is most frequently used in relation to written, typed or printed paper documents. Where information is stored by other means on other surfaces we accept that the storing item concerned is more readily referred to by reference to the means of storage or surface for storage concerned rather than as a 'document'. Hence reference to, for example, machines or tapes. However, terminological emphasis in description in such cases on the means or surface for recording information does not deprive

> such alternative stores of information from qualifying as 'documents' any more so than, for example, a tombstone, which is expressly included in the dictionary definition referred to. It seems to us that the essential essence of a document is that it is something concerning recorded information of some sort. It does not matter if, to be meaningful, the information requires to be processed in some way such as translation, decoding or electrical retrieval.[3]

1 Charles Hollander, *Documentary Evidence* (12th edn, Sweet & Maxwell 2015) para 7–22.
2 1997 JC 23, 1997 SLT 958 (HCJ).
3 1997 SLT 958, 960F-G.

Visual reading of a document

3.34 Although the meaning of 'document' has been construed widely, nevertheless it was held by the court in *Darby (Yvonne Beatrice) v DPP*[1] that a visual reading cannot be a document. This must be correct. Unless the reading is stored in some way that enables it to be read at a later date, the reading is merely a transitory phenomenon that can only be captured by a person who provides original testimony by giving evidence about his perception.[2]

1 [1995] RTR 294, (1995) 159 JP 533 (DC).
2 *Owen v Chesters* [1985] RTR 191 where a police officer gave evidence of the reading from a breath test machine; see also (this list is not exhaustive) *Denneny v Harding* [1986] RTR 350; *Mayon v DPP* [1988] RTR 281; *Greenaway v DPP* [1994] RTR 17, 158 JP 27 (DC).

3.35 But oral testimony may be provided in lieu of documentary evidence. In a number of breath specimen cases, the defendants' counsel have submitted that it is necessary to provide the print-out as documentary evidence of the output recorded by the machine, and that substitute evidence given by a police officer as to the machine output is not admissible.[1] In *Thom v DPP*,[2] the print-out from an Intoximeter was not produced, and the defence objected to testimony by a police officer as to what he had seen on the print-out. Clarke J addressed this point as follows:

> I can see no distinction in principle between evidence by a witness that he looked at his watch and read the time at, say, noon, and evidence from a witness that he looked at the Lion Intoximeter and that he read the proportion of alcohol in 100 millilitres of breath as being X.[3]

1 When radar speed meters were introduced in the late 1950s, police officers had to note down the reading in their notebooks, because this was the only method of recording a reading: J M W McBride, 'The radar speed meter' [1958] Crim LR 349.
2 [1994] RTR 11.
3 [1994] RTR 11, 14 G.

3.36 Likewise, in *Sneyd v DPP*,[1] when the print-out from an Intoximeter was not produced, the court accepted the police officer's testimony of what he had seen on the print-out provided by the device, rather than what he had seen on the screen. Rejecting the objection on the basis that the testimony was secondary evidence, Richards LJ held that 'it is well established that evidence both as to the results of the analysis and as to the reliability of the machine can be given either in the form of a written print-out or orally by the officer who carried out the procedure.'[2] He held that there was no difference between the oral evidence of the results shown on the print-out and oral

evidence of the results on the screen of the machine – both were not inadmissible hearsay. In *R (on the application of Leong) v DPP*,[3] Silber J applied the analysis of Richards LJ, holding admissible the oral evidence of the police officer's reading from a print out: 'Where, as in the present case, there is evidence that the machine is working properly, there is no reason why the police officer concerned cannot give admissible evidence of what he saw in the print-out.'[4]

1 [2006] EWHC 560 (Admin).
2 [2006] EWHC 560 (Admin), [32].
3 [2006] EWHC 1575 (Admin).
4 [2006] EWHC 1575 (Admin), [14].

Authentication

3.37 When a document is tendered as evidence of its contents, it is often accompanied by proof that the document 'has some specific connection to a person or organization, whether through authorship or some other relation'.[1] As noted by Austin J: 'Authentication is about showing that the document is what it is claimed to be, not about assessing, at the point of the adducing of the evidence, whether the document proves what the tendering party claims it proves.'[2] Similarly, where any object is tendered in evidence, an adequate foundation for admission will require testimony first that the object offered is the object which was involved in the incident, and further that the condition of the object is substantially unchanged.[3]

1 Kenneth S Broun (ed.), *McCormick on Evidence*, II (7th edn, West Publishing 2013), 83–85 [221].
2 *Australian Securities and Investment Commission v Rich* (2005) 216 ALR 320, [118], [2005] NSWSC 417.
3 Broun, *McCormick on Evidence* 13–16 [213].

3.38 Electronic evidence must also be authenticated, as for any other form of evidence. The authentication evidence for electronic evidence is even more critical,[1] and can occasionally be challenging.[2] Undoubtedly the use of technology has afforded us convenience and efficiency. But if parties and investigative authorities choose to use the fruits of technology, they must also accept the need to prove the authenticity and integrity of the evidence produced by technology, even though the cost of such proof might be considered to be high. This is particularly the case where authentication evidence will shed light on the latent assumptions and hidden errors inherent in electronic evidence, which could affect the accuracy of the electronic evidence itself.

1 Seng, 'Computer output as evidence' 159–66; Rosemary Pattenden, 'Authenticating "things" in English law: principles for adducing tangible evidence in common law jury trials' (2008) 12 E & P 290.
2 The challenge of proving that evidence in digital form is authentic was the subject of *R v Cochrane* [1993] Crim LR 48 (CA); see the chapter on authentication for a detailed discussion.

3.39 Authentication evidence may also demonstrate that the errors in question will not have an adverse effect on the evidence itself. For instance, in *DPP v McKeown; DPP v Jones*,[1] the clocks on the Intoximeter 3000 used to measure the breath alcohol values of the defendants were not accurate. For this reason, the defendants challenged the admissibility of the print-outs from the device. In addressing whether the accuracy of the clocks was relevant to the accuracy of the print-out readings, Lord Hoffmann examined the functioning of these devices and concluded that, for the purposes of s 69

of the Police and Criminal Evidence Act 1984,[2] a malfunction was irrelevant unless it affected the way in which the computer processes, stores or retrieves the information used to generate the statement.[3] On the facts, the clock was not part of the processing mechanism of the Intoximeter, and the convictions of the defendants based on the print-out readings were upheld.

1 [1997] 1 All ER 737, [1997] 2 Cr App R 155 (HL).

2 Section 69 of the Police and Criminal Evidence Act 1984 was repealed by s 60 of the Youth Justice and Criminal Evidence Act 1999, although the relevant case law remains useful authority.

3 [1997] 1 All ER 737, 744. A study later demonstrated that breath alcohol values measured on the Lion Intoximeter 3000 are not affected if the machine clock is incorrect by more than four minutes: R C Denny, 'The Intoximeter 3000 and the four minute fallacy' (1998) 38 Medicine, Science and the Law 163. Minor typographical errors on a print-out do not alter the validity of the results: *Reid v DPP*, The Times, 6 March 1998, 149 (QB).

3.40 This does not mean that authentication evidence will always have to be supplied for each item of evidence. In civil proceedings in England & Wales, a party is deemed to admit the authenticity of a document disclosed under the provisions of Civil Procedure Rule (CPR) 31 unless notice is served that the party wishes the document to be proved at trial. Thus where the authenticity of a document is questioned, the party raising the issue is required to do so at an early stage of the proceedings, thereby providing the party submitting the document the opportunity of gathering evidence to prove the veracity of the document.

3.41 See the chapter on authentication for a more detailed discussion.

Best evidence

3.42 The best evidence rule can be considered from two points of view. It can be regarded as an inclusionary rule under which whatever is the best evidence is admissible, thus overcoming exclusionary rules such as the hearsay rule; alternatively, it can be regarded as an exclusionary rule, so that anything which is not the best evidence is inadmissible. Since *Omychund v Barker*,[1] the majority of the cases have used the rule in an exclusionary way to deny the use of copies of documents when the absence of the original was not satisfactorily accounted for.

1 1 ATK 22, 49; 26 ER 15.

3.43 Reaction against this rule began in the nineteenth century,[1] and by the latter part of the twentieth century it was recognized that the best evidence rule was no longer as relevant as it once was. In *Kajala v Noble*,[2] Ackner LJ held that the rule is now confined to written documents in the strictest sense of the term. Echoing the robust comments of Lord Denning MR in *Garton v Hunter (Valuation Officer)*,[3] his Lordship said:

> The old rule, that a party must produce the best evidence that the nature of the case will allow, and that any less good evidence is to be excluded, has gone by the board long ago. The only remaining instance of it is that, if an original document is available in one's hands, one must produce it; that one cannot give secondary evidence by producing a copy. Nowadays we do not confine ourselves to the best evidence. We admit all relevant evidence. The goodness or badness of it goes

> only to weight, and not to admissibility .... In our judgment, the old rule is limited and confined to written documents in the strict sense of the term, and has no relevance to tapes or films.[4]

1 Malek, *Phipson on Evidence* para 7-42; see the discussion of Sargent J in the New Hampshire case of *Howley v Whipple* 48 N.H. 487 (1869) in respect of best evidence in the case of telegrams.
2 (1982) 75 Cr App R 149 (DC).
3 [1969] 2 QB 37, 44, [1969] 1 All ER 451, [1969] 2 WLR 86 (CA).
4 *Kajala v Noble* (1982) 75 Cr App R 149 (DC) 152; whether it is necessary to produce the original when a photocopy is adduced in evidence will depend upon whether the production of the original is relevant and necessary, for which see *Attorney-General v Lundin* (1982) 75 Cr App R 90.

3.44 By 1990, Lloyd LJ in *R v Governor Ex p Osman (No 1)* observed that the best evidence rule had become a rule of practice or procedure.[1] He also made the following remarks about the rule:

> … this court would be more than happy to say goodbye to the best evidence rule. We accept that it served an important purpose in the days of parchment and quill pens.[2] But since the invention of carbon paper and, still more, the photocopier and the telefacsimile machine, that purpose has largely gone. Where there is an allegation of forgery the court will obviously attach little, if any, weight to anything other than the original; so also if the copy produced in court is illegible. But to maintain a general exclusionary rule for these limited purposes is, in our view, hardly justifiable.[3]

1 *R v Governor Ex p Osman (No 1) sub nom Osman (No 1), Re* [1989] 3 All ER 701, [1990] 1 WLR 277 (DC).
2 It will be interesting to know how many ancient documents were previously admitted into evidence that were actually copies: *A Guide to Seals in the Public Record Office* (2nd edn, HMSO 1968) 30.
3 [1989] 3 All ER 701, [1990] 1 WLR 277 (DC), 308.

3.45 The best evidence rule has been effectively limited to requiring a party having possession of an original document who is relying on it for the statements recorded on the document (primary evidence) to not wilfully refuse to produce the original document as primary evidence, and instead produce copies or substitutes (secondary evidence) in its place.[1]

1 Colin Tapper, *Cross and Tapper on Evidence* (8th edn, Butterworths 1995) 748, ch XVIII, s 1: Proof of the Contents of a Document. A. The General Rule. 1. Statement and Illustrations of the General Rule. This statement of the rule was removed in subsequent editions. See also *Wayte (William Guy Alexander)* (1982) 76 Cr App R 110 (CA), where photostat copies of two letters were not admissible in circumstances where the party seeking to rely on the documents refused to produce the original letters.

3.46 Where good reasons exist for the failure to produce the original document, secondary evidence, even in the form of oral testimony, is permissible. This may be illustrated by the case of *Taylor v Chief Constable of Cheshire*,[1] a case involving the inadvertent destruction of evidence. In this case, video images of the accused allegedly committing theft from a store were recorded on the store video recorder, and the manager of the store, three police officers, and the lawyer for the accused later saw these recordings. When the case was heard, it transpired that new security officers had erased the recording of the video images. The magistrates permitted the witnesses to give evidence of what they saw on the video recording. An appeal was made that the best evidence – the video recording – could not be admitted because it had been

destroyed, and that testimonial evidence of the recording was not the best evidence. This was rejected. Although the best evidence in this instance was the video recording, the unavailability of this recording did not preclude the admission into evidence of the testimony of those witnesses who viewed the recording. The recollections of the witnesses ought not be precluded because the best evidence was not available. The evidence offered by the witnesses was, as pointed out by Ralph Gibson LJ, 'direct evidence of what was seen to be happening in a particular place at a particular time', and it was for the trier of the facts to assess its weight, credibility and reliability.[2]

1 [1987] 1 All ER 225, [1986] 1 WLR 1479 (QB).
2 [1987] 1 All ER 225, 230.

3.47 Since the statutory intercession of the Civil Evidence Act 1995 and the Criminal Justice Act 2003, the best evidence rule has further taken a simplified, statutory form. The judgment of the Court of Appeal in *Masquerade Music Ltd v Springsteen*,[1] suggests that the best evidence rule is hardly of any relevance. After considering the best evidence rule in detail and reviewing the case law extensively,[2] Jonathan Parker LJ outlined the position with respect to the best evidence rule in the twenty-first century:

> In my judgment, the time has now come when it can be said with confidence that the best evidence rule, long on its deathbed, has finally expired. In every case where a party seeks to adduce secondary evidence of the contents of a document, it is a matter for the court to decide, in the light of all the circumstances of the case, what (if any) weight to attach to that evidence. Where the party seeking to adduce the secondary evidence could readily produce the document, it may be expected that (absent some special circumstances) the court will decline to admit the secondary evidence on the ground that it is worthless. At the other extreme, where the party seeking to adduce the secondary evidence genuinely cannot produce the document, it may be expected that (absent some special circumstances) the court will admit the secondary evidence and attach such weight to it as it considers appropriate in all the circumstances. In cases falling between those two extremes, it is for the court to make a judgment as to whether in all the circumstances any weight should be attached to the secondary evidence. Thus, the 'admissibility' of secondary evidence of the contents of documents is, in my judgment, entirely dependent upon whether or not any weight is to be attached to that evidence. And whether or not any weight is to be attached to such secondary evidence is a matter for the court to decide, taking into account all the circumstances of the particular case.[3]

1 [2001] EWCA Civ 513, [2001] EMLR 654, [2001] All ER (D) 101 (Apr).
2 [2001] EMLR 654, [64]–[85].
3 [2001] EMLR 654, [85].

3.48 Waller and Laws LJJ concurred. In other words, there is no automatic bar to the failure to admit the original document as primary evidence. Instead, when the original document is no longer available, a copy of the original evidence is admissible but an adjudicator must consider its weight as secondary evidence.

3.49 The modern application of this rule is illustrated by *Post Office Counters Ltd v Mahida*.[1] In this case, the Post Office sought to claim an alleged deficiency of social security benefits paid out against the defendant, the sub-postmaster general. The deficiency was set out in a schedule prepared by investigators of the Post Office based on checks conducted against the underlying dockets and foils. Subsequently the

dockets and foils were destroyed as part of a routine process. The trial judge accepted the schedule as secondary evidence and found against the defendant. On appeal, the Court of Appeal was concerned that the secondary evidence was of insufficient weight to prove the precise amount of the debt claimed against the defendant. In particular, the Post Office as an institution could not readily be said to have discharged the burden of proving the precise amount of the debt when it was alleged that the defendant had been responsible for this loss, and denied the defendant the opportunity to check those figures.[2] For this reason, the very basic unfairness should have led the trial judge to consider that the amount of the debt was not proved, and the defendant's appeal was allowed.

1 [2003] EWCA Civ 1583.
2 [2003] EWCA Civ 1583, [27].

Analogue evidence

3.50 Although the best evidence rule is now tightly confined, it applies to both civil and criminal proceedings.[1] But as the statutory formulations of the rule in s 8 of the Civil Evidence Act 1995 and s 133 of the Criminal Justice Act 2003 retain the difference between primary and secondary evidence, the ramifications are different, depending on whether the evidence is in analogue or in electronic form.

1 *R v Wayte* (1982) 76 Cr App R 110.

3.51 In the physical world, the primary evidence is an original document, and the secondary evidence is in the form of copies of the original. The best evidence rule will require the production of the original document to prove the content in question, and the submission of copies is considered inferior evidence. But the fact that copies were made, for instance, by a reprographic process such as photocopying, will not prevent the copies themselves from being originals. In *Miller-Foulds v Secretary of State for Constitutional Affairs*[1] regarding orders issued by Brentford County Court, Pelling J noted the following:

> The method of production involved copying an original draft [order] and then sealing the copies thus resulting. The copies, once sealed, were original orders. The original draft was just that: a draft. The fact that the documents that were sealed were produced by photocopying rather than copying out by hand the same document umpteen times is wholly irrelevant, in my judgment, as long as the document itself resulting from the copying process was sealed.[2]

1 [2008] EWHC 3443 (Ch). A subsequent application before Lloyd LJ was rejected, for which see *Miller-Foulds v Secretary of State for Justice* [2009] EWCA Civ 1132.
2 [2008] EWHC 3443 (Ch), [26].

3.52 The concepts of 'primary' and 'secondary' evidence take a different shape when applied to material objects that must be processed to be viewed. Consider, for instance, a photograph taken with a camera containing film, or a plate. The negative or the plate comprises the only copy of the image in reverse.[1] It is the negative or plate that is the material upon which the primary evidence is recorded. However, few people will be satisfied by looking at the primary image, if only because it is not easy to view, and is not intended to be viewed in this form, unless by means of a projector (if the primary

image is a negative). This means that the printed image is secondary evidence. Any number of copies of the primary object can be made, although no printed copy will be an exact copy of the film or plate. This is because the processes applied and the mix of chemicals used in transforming the negative into a print will determine how accurately the photograph reflects the image, in particular the degree of contrast (that is the range of grey tones) captured on the negative. For example, the degree of contrast will affect how bruising is reproduced on the photograph: a high contrast makes the bruising appear darker and more dramatic, while a low contrast will lessen the effect of the visual image, making the bruise seem somewhat less consequential.

1 A point noted by Smith LJ in *Griffiths v DPP* [2007] RTR 44, [21].

Digital evidence

3.53 In contrast to the discussion above, the range of evidence in digital form is vast, and it comprises not just print-outs of what might be termed conventional files, such as copies of letters, contracts or spreadsheets. Other forms of digital documents include reports from computer databases, the electronic records of transactions and the digital store and reproduction of images, such as the scanned image of an original paper document. The treatment of evidence in digital form calls for different and occasionally difficult considerations.

3.54 First, there may be issues identifying the primary evidence of a digital document. In *Derby v Weldon (No. 9)*, Vinelott J considered the memory or database of a word-processor or computer to be the 'original document',[1] presumably on the basis that these are components 'on which material fed into a simple word processor is stored'.[2] However, Professor Tapper disagrees, and takes the view that the print-out from the word-processed electronic document is the original and the document in memory computer is the copy.[3] Both views are possible. Vinelott J's analysis is plausible – where the print-out is generated as a physical draft to aid in the editing of the word-processed document. But Professor Tapper's view could also be justified where the object behind the use of the word-processor is the generation of the print-out as the final, definitive version of the document. In such a case, the authentic print-out may be a better form of evidence than the state of the document in internal memory at a later time. This inversion provides a good illustration of the danger of assuming that the print-out may not be the best evidence in any given situation.

1 *Derby v Weldon (No. 9)* [1991] 2 All ER 901, 906.
2 *Derby v Weldon (No. 9)* [1991] 2 All ER 901, 906.
3 Colin Tapper, 'Evanescent evidence' (1993) 1 Intl J L & Info Tech 35, 42.

3.55 In addition, the use of a digital device need not always produce an 'original document'. Where the 'original document' is created in digital form but is never stored in a more permanent, non-ephemeral manner, the 'original' digital 'document' ceases to exist for all practical purposes. Instant messaging is an example of evidence that might not be stored, which makes it analogous to an oral conversation.

3.56 The issues may be further considered with the following extended illustration. For instance, the original of a physical document, such as a commercial contract between two parties, signed by the authorized representatives of both parties and

acknowledged as the original, is primary evidence of the content of the contract. Even if the contract was created on a computer, the physical document will still be the original document as it was signed and adopted by both parties.[1] However, should the contract, which is subsequently acted upon by both parties, only exist in digital form on a computer, the primary evidence of the document will be the digital contract residing on an identified computer storage device such as the hard drive of a computer. Printing the document out on paper will provide copies in a human-readable form, which will in turn comprise secondary evidence of the document.[2]

1 The physical document might have a digital counterpart, as in Austria, for which see Friedrich Schwank, 'CyberDOC and e-Government: the electronic archive of Austrian notaries' (2004) 1 Digital Evidence and Signature Law Review 30, 32.

2 The schedule produced in *R v Nazeer* [1998] Crim LR 750 cannot be considered to be hearsay or secondary evidence, because it was real evidence produced by individuals using different sources of information (including computer records).

3.57 Now consider the matter one stage further. Assume the original digital file is accessed multiple times after the contract is executed, but its file contents are not altered: perhaps particular clauses are copied for other reasons. The metadata for the digital file may have been changed to record the action of opening and closing the file, even if no substantive changes are made. Although the metadata might have been altered, the content of the file in question has not been affected. In these circumstances, it might be considered that the integrity of the original digital data is compromised. But as the content (rather than the metadata) of the digital document is unchanged, the digital document remains the primary evidence, and a print-out of that document is a faithful copy of the original.[1] The metadata can be compared to a file register in the physical world that records the name of the person to whom the physical file was given, the date and time the person obtained the file, and the date and time it was returned: the register information does not alter the content of the statements made in the file (unless the person obtaining access to the file alters its contents). In such circumstances, the metadata does not affect the integrity of the digital data, which makes the secondary evidence of the file in the form of the print-out a reliable reproduction of the digital file.

1 Professor Tapper expressed the contrary view, that 'the memory holds the copy and the original is the printed copy', in Tapper, *Cross and Tapper on Evidence* 35, 42. This is correct if the printed version is a document such as a contract, where the contract is subsequently signed by the parties with manuscript signatures and excludes reference to any other version.

3.58 Consider another example: the drafting of a contract by an external lawyer for a multinational company. The task will comprise a number of stages, including liaising with a number of people internally with different responsibilities to produce an initial draft of the contract; it will be passed to the other contracting party for its comments, before, after a substantial period of negotiation, a final version is produced to the satisfaction of both parties. In all probability, various versions of the draft contract will exist in storage devices on computers, hand-held devices and back-up devices belonging to several companies and their employees, perhaps across different jurisdictions. If the contract is then printed and signed by the authorized representatives of the two parties, the original document will be the printed version. If the issue is as regards a particular version of the contract at a particular point in the negotiations, the draft digital version of the contract will be original evidence because that electronic copy

is the best evidence of that version of the contract and a print-out of that version is secondary evidence.

3.59 In addition, digital documents may themselves be stored, changed, compiled and collected into new documents, and the new documents may be original documents in themselves. The Canadian case of *R v Bell*[1] is instructional in this regard. In this case, the bank's computer software processed the various transactions of its customers' chequing accounts into a monthly statement for each account. Two identical copies of the monthly statement were printed, one for the customer, and one for the bank. The bank retained its copy of the monthly statement, but did not retain a record of the transactions. The trial judge held that a copy of the statement was not admissible because the transaction information stored on a computer was the record, and the original 'record' as a record of a financial institution (and its subsequent copy) no longer existed. On appeal, this analysis was rejected. Weatherston JA noted that the form in which information is recorded may change from time to time, and a new form in which information is recorded, such as a compilation or collection of other records, is equally a record of that kind of information. The court found the monthly statement to be such a 'record' that consolidated the transactions of a financial institution and allowed the appeal.[2]

1 (1982) 35 OR (2d) 164 (CA).
2 (1982) 35 OR (2d) 164 (CA), [13].

Civil proceedings

3.60 The admissibility of secondary evidence in civil proceedings is governed by of s 8 of the Civil Evidence Act 1995, which permits the introduction of copies of documents into evidence for the purpose of proving the statement contained in the document:

> 8.—(1) Where a statement contained in a document is admissible as evidence in civil proceedings, it may be proved—
>
> (a) by the production of that document, or
>
> (b) whether or not that document is still in existence, by the production of a copy of that document or of the material part of it,
>
> authenticated in such manner as the court may approve.
>
> (2) It is immaterial for this purpose how many removes there are between a copy and the original.

3.61 A 'document' is in turn defined in s 13 as 'anything in which information of any description is recorded', and 'copy' of a document as 'anything onto which information recorded in the document has been copied, by whatever means and whether directly or indirectly'. There are two operative parts to s 8. Section 8(1)(a) provides that an admissible statement contained in a document may be proved by the production of the original document itself. Section 8(1)(b) provides that the same document may be proved by the production of a copy of that document or a material part of it, with the expression 'whether or not that [primary] document is still in existence' completely eviscerating the common law best evidence rule. And although s 8(1) uses the language of 'a statement contained in a document', suggesting that the statutory version of the best evidence rule only applies to documentary evidence used in a testimonial sense, a better reading is that s 8 applies to both documentary evidence as testimonial

evidence and documentary evidence as real evidence.[1] This means that s 8 will apply to the analogue record of the measurements of a device (the measurement constitutes the statement of the document)[2] or the print-out from an Intoximeter.

1 Tapper, *Cross and Tapper on Evidence* 669.
2 Such as the film in *The Statute of Liberty, Sapporo Maru M/S (Owners) v Steam Tanker Statute of Liberty (Owners)* [1968] 2 All ER 195.

3.62 The admissibility of the copied document as secondary evidence is subject to one condition and one qualification. The condition is that, as set out in the proviso to s 8(1), the copied document must be 'authenticated in such manner as the court may approve', just as the primary document must be authenticated. In other words, where the credibility of the digital data is in question, foundation evidence, typically in the form of testimony, will have to be introduced and tested to determine whether the secondary evidence can be accepted as 'a copy' of the original document. The residual judicial control over the admissibility of secondary evidence takes the form of judicial prescription of the requisite authentication evidence to prove that it is an accurate and reliable copy of the whole or a material part of the original document.

3.63 The qualification is that, by s 8(2), the number of removes between the copy and the original document is statutorily deemed to be irrelevant. This detracts from the judicial control role as explained above, and also undermines the judicial assessment of the authentication evidence as to the true accuracy and reliability of the secondary evidence.

Criminal proceedings

3.64 The starting point for the application of the best evidence rule in criminal proceedings is s 133 of the Criminal Justice Act 2003:

> 133 Proof of statements in documents
>
> Where a statement in a document is admissible as evidence in criminal proceedings, the statement may be proved by producing either-
>
> (a) the document, or
>
> (b) (whether or not the document exists) a copy of the document or of the material part of it,
>
> authenticated in whatever way the court may approve.

3.65 The s 133 provisions are identical to those for civil proceedings in the Civil Evidence Act 1995, save for the fact that there is no longer a mention of the number of times a copy is removed from the original in s 133 in the Criminal Justice Act. (It is suggested that the elimination of the number of removes qualification in s 133 is an improvement over the equivalent formulation of the best evidence rule in the Civil Evidence Act, in removing getting rid of the judicial handicap for assessment of the authentication evidence.) The other difference is that proof in criminal proceedings must rise to the appropriate standard, which is proof beyond reasonable doubt in the case of the prosecution, and proof on the balance of probabilities in the case of the defence.[1] Otherwise, it should also be noted that notwithstanding the reference to 'a statement in a document', for the same reasons as outlined above in relation to the Civil

Evidence Act 1995, the best evidence provisions should apply equally to a document as real evidence as to a document as testimonial evidence.[2] In other words, as in civil proceedings, secondary evidence of an electronic document is admissible subject to authentication evidence.

1 Tapper, *Cross and Tapper on Evidence* 610.
2 Note that in *R v Minors & Harper* (1989) 89 Cr App R 102, it was held that s 24, Criminal Justice Act 1988 only applied to a 'statement in a document' and not to real evidence. s 24, like s 27, the predecessor provision to s 133, is found in Part II (Documentary Evidence in Criminal Proceedings) of the Criminal Justice Act 1988. That notwithstanding, it could be argued that the holding in *R v Minors & Harper* should be confined to s 24 (an exception to the hearsay rule), and has no application to the interpretation of s 27 (a restatement of the best evidence rule).

3.66 The effect is that while the original electronic document, if available, should be adduced into evidence, in practice, a copy of the document tends to be adduced as secondary evidence. The copy may be at least one, if not two, removes[1] from the original. This should not matter, provided the digital copy has been copied in a way that captures the file in its entirety, including all its attributes, such as the metadata, without altering the original data. (On this point, please see the detailed discussion in the chapter dealing with authentication.)

1 It is usually two removes from the original, if the original is considered to be the operational electronic document that is actively used on the computer system in question, and a copy is previously taken from that operational electronic document (in computer science terms, a 'snapshot'– the state of the system at a particular point in time, considering that some time would have lapsed between the taking of this copy and the currently operational version of the electronic document), and a copy is in turn taken from that previous copy for purposes of preparation of proceedings.

3.67 To a certain extent, rather than question whether a document in digital form is an original or a copy, it might be more useful and relevant to refer to the proof of authenticity, or provenance, or reliability of a digital file. Such is required under both s 133 of the Criminal Justice Act 2003 as well as s 8 of the Civil Evidence Act 1995. This in turn encapsulates proof of the integrity of the content of the data. Because of the ease in which a digital document may be migrated from one storage device to another, and undergo format and other changes, including content and metadata changes, it is vital to require any such changes to be documented in such a way as to preserve the integrity and authenticity of the copy. Thus it might be more relevant, when referring to digital data, to concentrate on establishing which version of the data is required, particularly whether the making of copies of the digital document is properly documented.

Admissibility

3.68 Evidence is admitted into legal proceedings if it is relevant to an issue in dispute, subject to a number of exceptions.[1] It is a matter of law for a judge to determine whether evidence is admissible. Generally, judges are required to determine whether evidence is to be excluded in criminal trials far more frequently than in civil matters, especially where admitting the evidence might not be in the interests of justice.[2] For instance, in *R v Fowden and White*[3] the Court of Appeal held that a video film showing activities that were consistent with the acts of theft had been improperly admitted. The prejudicial value outweighed its probative effect, because the witnesses that identified the accused knew them from a similar case of theft that occurred a week

after the events recorded in the video film, and the defence was therefore not able to test the accuracy of the identification without causing prejudice and embarrassment.[4]

1 For a more detailed discussion, see Malek, *Phipson on Evidence*, ch 2 and paras 7-01 to 7-16. For a brief consideration of a number of jurisdictions, see Olivier Leroux, 'Legal admissibility of electronic evidence' (2004) 18 Intl Review L Computers & Tech 193.
2 Police and Criminal Evidence Act 1984, s 78; Criminal Justice Act 2003, s 114(1)(d).
3 [1982] Crim LR 588.
4 In *R v Caldwell, R v Dixon* (1993) 99 Cr App R 73, 78 the members of the court considered it would be useful to have a set of procedures in relation to the use of video recordings for the purposes of identification.

3.69 In civil proceedings, evidence that is admissible can be excluded in accordance with the provisions of CPR 32.1(2), which provides a judge with the explicit general power to exclude evidence when in the role of managing a case:

> 32.1 (1) The court may control the evidence by giving directions as to –
>
> (a) the issues on which it requires evidence;
>
> (b) the nature of the evidence which it requires to decide those issues; and
>
> (c) the way in which the evidence is to be placed before the court.
>
> (2) The court may use its power under this rule to exclude evidence that would otherwise be admissible.

3.70 However, the power, as pointed out by Arden LJ, in adopting the argument of the appellants in *Great Future International Ltd v Sealand Housing Corporation*, 'must be used with great circumspection for the purpose of achieving the overriding objective.'[1] Professor Tapper notes that the modern tendency is to admit evidence, and then consider its weight,[2] as illustrated by the comment of Cockburn CJ in *The Queen v Churchwardens, Overseers and Guardians of the Poor of the Parish of Birmingham*: 'People were formerly frightened out of their wits about admitting evidence lest juries should go wrong. In modern times we admit the evidence and discuss its weight.'[3]

1 [2002] EWCA Civ 1183, [24].
2 Tapper, *Cross and Tapper on Evidence* 74.
3 (1861) 1 B & S 763, 767; 121 ER 897.

Weight

3.71 The questions of weight, credibility and sufficiency of the evidence are decisions for the members of a jury, and for the judge where a case is tried without a jury. There are no fixed rules to determine what weight to give to any item of evidence. In *R v Madhub Chunder Giri Mohunt*, Birch J observed: 'For weighing evidence and drawing inferences from it, there can be no canon. Each case represents its own peculiarities and in each common sense and shrewdness must be brought to bear upon the facts elicited'[1] and Lord Blackburn commented in *Lord Advocate v Blantyre* that 'The weight of evidence depends on rules of common sense.'[2]

1 (1874) 21 W.R.Cr (India) 13, 19.
2 (1879) 4 App Cas 770, 792.

3.72 When conducting a trial with members of a jury, the judge may withdraw an issue because the proponent has failed to adduce sufficient evidence in support of the claim. Furthermore, in summing up to the members of the jury at the end of the trial, the judge is required to provide directions on a range of issues, including, but not limited to: who has the burden of proof; what presumptions, if any apply; when supporting evidence should be considered before putting weight on certain types of evidence; and to offer comments on matters including the weight of the evidence, although it must be made explicit that such comments are meant to help the members of the jury, because they must reach their own decision.[1] In addition, there are a number of factors set out in s 114(2) of the Criminal Justice Act 2003 that deal with the assessment of weight of hearsay in criminal proceedings.

1 *The Crown Court Bench Book: Directing the Jury* issued by the Judicial Studies Board was available online at <www.judiciary.gov.uk/wp-content/uploads/JCO/Documents/Training/benchbook_criminal_2010.pdf>.

Execution and electronic signatures

3.73 Public documents such as birth and death registers, registers of baptisms and marriages, Acts of Parliament, royal proclamations, Orders in Council, statutory instruments and journals of either House of Parliament, may be proved in evidence by the mere production of the appropriate copy, certified or seal where appropriate. Proof of their execution is also dispensed with.[1] But the court requires proof of the due execution of a private document, unless it is more than 20 years old and comes from the proper custody.[2] 'Due execution of a private document is proved by showing it was signed by the person by whom it purports to have been signed, and, where necessary, attested.'[3] Out of this, a substantial body of case law has arisen to guide the proof of physical signatures. But despite the early acceptance of electronic evidence in case law, until recently, signatures as applied to electronic documents were operating in the shadows of common law rules relating to physical signatures.[4]

1 Tapper, *Cross and Tapper on Evidence* 669–74.
2 Tapper, *Cross and Tapper on Evidence* 674.
3 Tapper, *Cross and Tapper on Evidence* 674.
4 Stephen Mason, *Electronic Signatures in Law* (4th edn, University of London 2016) (the strength of this text lies in the extensive case law); Lorna Brazell, *Electronic Signatures and Identities Law and Regulation* (2nd edn, Sweet & Maxwell 2008) (the strength of this text lies in the regulatory framework for digital signatures discussed); George Dimitrov, *Liability of Certification Service Providers* (VDM Verlag Dr. Müller 2008); M H M Schellenkens, *Electronic Signatures Authentication Technology from a Legal Perspective* (TCM Asser Press 2004); Dennis Campbell (ed.), *E-Commerce and the Law of Digital Signatures* (Oceana Publications 2005). For translations of electronic signature cases from across the world into English, see also the Digital Evidence and Electronic Signature Law Review.

3.74 The Electronic Communications Act 2000, which extends to Northern Ireland, received the Royal Assent on 25 May 2000,[1] and was amended in 2016 by The Electronic Identification and Trust Services for Electronic Transactions Regulations 2016 (SI 2016 No 696).[2] The amended definition of an electronic signature reads in s 7(2) as follows:

> (2) For the purposes of this section an electronic signature is so much of anything in electronic form as-

> (a) is incorporated into or otherwise logically associated with any electronic communication or electronic data; and
>
> (b) purports to be used by the individual creating it to sign.

1 Regulation of Investigatory Powers Act 2000, s 16 (5).
2 Made on 30 June 2016, laid before Parliament 1 July 2016, into force on 22 July 2016, implementing the changes brought about by Regulation (EU) No 910/2014 of the European Parliament and of the Council of 23 July 2014 on electronic identification and trust services for electronic transactions in the internal market and repealing Directive 1999/93/EC, OJ L257, 28.8.2014, 73–114.

3.75 Section 7(1) of the Act provides for the admissibility of the electronic signature in two ways:

> 7(1) In any legal proceedings-
>
> (a) an electronic signature incorporated into or logically associated with a particular electronic communication or particular electronic data, and
>
> (b) the certification by any person of such a signature,
>
> shall each be admissible in evidence in relation to any question as to the authenticity of the communication or data or as to the integrity of the communication or data.

3.76 An electronic signature is admissible under the provisions of s 7(1)(a) where it is incorporated into or logically associated with a particular electronic communication or data. Alternatively, in accordance with the provisions of s 7(1)(b), the certification by any person of such an electronic signature is admissible as to the authenticity or the integrity of the communication or data. The certificate would normally be provided by an entity such as a trusted third party, although the provision does not rule out self-certification.

3.77 There are various types of signatures, all of which can demonstrate the intent of the signing party to authenticate the document. For physical signatures, the act of the person writing a manuscript signature or applying the impression of the seal is the act of intent, and the evidence of the act is the physical manifestation of the signature by the application of ink on to paper, or the wax placed onto the surface of the material. In the same way, a signature in electronic form is the act of the person doing an act or series of acts, which may comprise more than one act at different times, which is subsequently manifested in human-readable form.

3.78 The following are some of the different types of electronic signatures that are recognized:[1] typing a name into a document;[2] an email address;[3] clicking the 'I accept' icon; a PIN;[4] biodynamic signature; scanned manuscript signature, and a digital signature.[5]

1 See Mason, *Electronic Signatures in Law*, for a detailed survey of the different forms of electronic signature and case law.
2 *Hall v Cognos Limited* (Hull Industrial Tribunal, 1997) Case No 1803325/97.
3 In *J Pereira Fernandes SA v Mehta* [[2006] EWHC 813 (Ch); [2006] 1 WLR 1543; [2006] 2 All ER 891; [2006] 1 All ER (Comm) 885; [2006] All ER (D) 264 (Apr); [2006] IP & T 546; (2006) The Times 16 May 18 (in respect of the Statute of Frauds 1677, s 4), the judge reached a conclusion that is difficult to reconcile with the international cases or long-standing English case law – for a comprehensive analysis, see Mason, *Electronic Signatures in Law* paras 11.4–11.41.
4 The banks have led the way in the use of PINs, and now rely on technology to a great extent. For a PIN case in England, see *Job v Halifax PLC* (2009) (Case No 7BQ00307): the judgment is published in

full in (2009) 6 Digital Evidence and Electronic Signature Law Review 235, 245, with a commentary by Alistair Kelman.

5 As pointed out by Ugo Bechini, a 'manuscript signature links a document to a person, while a digital signature does not: it links a document to a device' (Ugo Bechini, 'Bread and donkey for breakfast. How IT law false friends can confound lawmakers: an Italian tale about digital signatures' (2009) 6 Digital Evidence and Electronic Signature Law Review 80.

3.79 The fact that an electronic signature is used to authenticate an electronic document and establish its integrity does not absolve the party who has the burden of proving the document from authenticating the electronic signature itself. This may call for the submission of evidence such as extrinsic evidence to demonstrate that the electronic signature establishes one or more aspects of the authenticity or integrity or both of the electronic document as set out in s 15(2) of the Electronic Communications Act 2000.

Video and audio evidence

Testimonial use in legal proceedings

3.80 In exceptional instances, video-recorded and tape-recorded evidence may be used in lieu of testimonial evidence. In civil proceedings, evidence may be given by means of a video link or any other means, subject to leave being obtained from the court.[1] In criminal matters, it is possible to record the initial interview with children,[2] and admit the recording in evidence, subject to leave of the court and any editing that the court decides is necessary.[3] Leave is required to adduce a video recording of the testimony of a witness in accordance with the provisions of s 27 of the Youth Justice and Criminal Evidence Act 1999.[4]

1 CPR 32.3, which is supplemented by Practice Direction 32 – Evidence Annex 3. See also the *Admiralty and Commercial Courts Guide*, app 14 and the *Chancery Court Guide*, ch 14.

2 Section 35A of the Criminal Justice Act 1988 was added by s 54 of the Criminal Justice Act 1991.

3 Criminal Justice Act 1988, s 35A(2).

4 For further details, see James Richardson (ed), *Archbold: Criminal Pleading, Evidence and Practice* (65th rev edn, Sweet & Maxwell 2017); David Ormerod and David Perry (eds), *Blackstone's Criminal Practice* (Oxford University Press 2017); Barbara Barnes (ed), *Archbold: Magistrates' Courts Criminal Practice* 2017 (14th rev edn, Sweet & Maxwell 2017).

3.81 Video-conferencing and web-conferencing technology has also made it possible to provide testimonial evidence outside the court.

Identification and recognition evidence

3.82 Surveillance cameras are very much part of life in the twenty-first century, ever since the foundations of their use were laid in the latter decades of the twentieth century. Evidence of images from security cameras can be very helpful in identifying the perpetrators of crimes. Such evidence has been admitted in English courts, mainly in criminal cases.[1] The widespread availability of video-recorded and tape-recorded evidence has opened up the possibility that such evidence may be augmented with more advanced techniques, and the enhancement of the sounds or images, together with the use of more advanced techniques such as aural identification and facial mapping, can help to identify the parties in a recording.

1 A list that is not exhaustive includes: *McShane* (1978) 66 Cr App R 97; *R v Fowden and White* [1982] Crim LR 588 (CA); *R v Grimer* [1982] Crim LR 674, 126 SJ 641 (CA); *R v Dodson (Patrick); R v Williams (Danny Fitzalbert Williams)* [1984] 1 WLR 971, (1984) 79 Cr App R 220; *Stockwell (Christopher James)* (1993) 97 Cr App R 260; *Clarke (Robert Lee)* [1995] 2 Cr App R 425; *Clare (Richard), Peach (Nicholas William)* [1995] 2 Cr App R 333; *R v Feltis (Jeremy)* [1996] EWCA Crim 776; *R v Hookway* [1999] Crim LR 750; *R v Briddick* [2001] EWCA Crim 984; *Loveridge (William)* [2001] EWCA Crim 973, [2001] 2 Cr App R 29. In this instance, the accused were recorded by video in the court, an act which was prohibited by s 41 of the Criminal Justice Act 1925, and the recording was also held to have infringed the rights of the accused under art 8 of the Human Rights Act 1998 – however, neither infringement was held to have interfered with the right to a fair trial (E. Goldstein, 'Photographic and videotape evidence in the criminal courts of England and Canada' [1987] Crim LR 384).

3.83 Before such evidence is used, there should be a careful examination[1] of the technology in question. A good example of this judicial scrutiny is that done by Steyn LJ in *Clarke (Robert Lee)*,[2] where his Lordship analysed the technique of facial mapping[3] by video superimposition. The court carefully considered the reliability of the underlying scientific techniques, noting that the techniques themselves could be fit for debate, and their improper use by an expert in the particular case could in turn affect the probative value of such evidence. It was only after it was satisfied on these two grounds that the identification evidence from the application of the technique was admitted.

1 The careful examination may be done in a trial within a trial, also called a 'voir dire'.
2 [1995] 2 Cr App R 425, 430F.
3 Michael C Bromby, 'At face value?' (2003) NLJ Expert Witness Supplement 301, 302–4; *R v Jung* [2006] NSWSC 658.

3.84 Issues regarding the reliability and application of these techniques are very much for expert evidence, depending on the nature and sophistication of each technique. But some guidance may be sought that stem from the best practices for handling electronic evidence. For instance, for evidential techniques that involve manipulating and enhancing digital imagery, Gregory Joseph has noted that the following steps must be taken before enhanced digital imagery can usefully be used:[1]

> 1. The original image needs to be properly authenticated.
>
> 2. The original image must remain intact to enable the original to be compared with the enhanced version.
>
> 3. The original image should be preserved in such a way that its integrity cannot be impugned.
>
> 4. The process of enhancement should be fully documented.
>
> 5. The process of enhancement should be carried out in such a way that the process can be repeated by the other party.
>
> 6. The enhanced images should be preserved in such a way that prevents it from being manipulated and thereby preserves its integrity.

1 Gregory P Joseph, *Modern Visual Evidence* (Law Journal Press 2009) 4.

3.85 Important lessons were also spelt out regarding the use of voice recognition technologies and techniques for identification purposes in *R v Flynn and St John*.[1] In this case, the prosecution sought to identify the two appellants as conspirators of a robbery through voice recognition techniques. Before the robbery, the police secretly fitted a listening and transmitting device to one of the vehicles it was assumed (correctly) that the conspirators would use for the robbery. Four police officers testified that they

recognized the appellants' voices from the 60 minutes of covert recording made by the device. The trial judge ruled admissible the evidence of the police officers and the transcripts of the recording and placed the evidence before the jury. The appellants challenged the decision of the trial judge to admit the voice recognition evidence of the officers and the judge's failure to give an appropriate direction to this evidence.

1 [2008] EWCA Crim 970, [2008] 2 Cr App R 20, [2008] Crim LR 799.

3.86 In giving judgment on appeal, Gage LJ noted that there are two categories of voice recognition evidence: expert evidence using either auditory analysis or acoustic/spectrographic analysis, or lay listener evidence, where the lay listener as a witness is required to possess some special knowledge of the suspect that enables him to recognize the suspect's voice. Such witnesses may be close relatives or friends, but they may also be persons who acquire such familiarity by the frequency of their contact with the suspect. Gage LJ also noted that suspect identification by voice recognition is more difficult than visual identification, that voice identification by experts using sophisticated auditory, acoustic and spectrographic and that sophisticated auditory techniques is likely to be more reliable than identification by a lay listener, and that the quality of identification by a lay listener is highly variable. In addition, research has shown that a confident recognition by a lay listener of a familiar voice may nevertheless be wrong, because while an expert is able to draw up an overall profile of the individual's speech patterns, in combination with instrumental analysis and reference research, a lay listener's response is fundamentally opaque because he cannot know and has no way of explaining which aspects of the speaker's speech patterns he is responding to, and has no way of assessing the significance of the individually observed features relative to the overall speech profile. This makes it more difficult to challenge the accuracy of his evidence.

3.87 For all these reasons, the Court of Appeal allowed the appeal, holding that the police officers as lay listeners had a limited opportunity to acquire familiarity with the appellants' voices, and that the quality of the covert recording was poor. In contrast, both experts, one representing the prosecution and the other representing the appellants, were unable to recognize their voices, further casting doubt on the officers' voice recognition evidence.

3.88 While *R v Flynn and St John* did not close the door on voice recognition evidence, in a paper by Gary Edmond, Kristy Martire and Mehera San Roquem, the authors suggest the following minimal safeguards required before the prosecution can seek to admit voice recognition evidence from lay listeners:

> 1. The process must be properly recorded, and the amount of time spent in contact with the defendant will be very relevant to the issue of familiarity.
>
> 2. The date and time spent by the police officer compiling a transcript of a covert recording must be recorded. If the police officer annotates the transcript with his views as to which person is speaking, that must be noted.
>
> 3. A police officer attempting the voice recognition exercise must do so without the aid of a transcript that bears another officer's annotations of whom he believes is speaking.
>
> 4. It is highly desirable that a voice recognition exercise should be carried out by someone other than an officer investigating the offence.[1]

1 [2008] EWCA Crim 970, [53]; also the paper by Gary Edmond, Kristy Martire and Mehara San Roque, 'Unsound law: issues with ("expert") voice comparison evidence' (2011) 35 Melbourne University Law Review 52.

3.89 These safeguards are certainly in line with the issues raised by Gage LJ in *R v Flynn and St John*, and highlight the care with which both the parties and the courts must observe when seeking to admit computer-generated and computer-augmented evidence, in order to safeguard the evidential process.

Computer generated animations and simulations

3.90 Digital visual evidence presentation systems (including digital displays, computer-generated graphical presentations, animated graphics and immersive virtual environment technology) have been used in legal proceedings in many jurisdictions. Such tools can be used to present evidence and illustrate hypotheses based on scientific data, or to depict the perception of a witness, or to illustrate what may have occurred (seen from a specific viewpoint) during a particular incident. Digital reconstruction technology may also be applied in a court to explore and illustrate 'what if' scenarios and questions, to test competing hypotheses and to expose any possible inconsistencies and discrepancies within the evidence.

3.91 Computer animations and interactive virtual simulations are potentially unparalleled in their capabilities for presenting complex evidence.[1] The use of such enabling visualization technologies can affect the manner in which evidence is assimilated and correlated by the viewer. In many instances, visual media can potentially help make the evidence more relevant and easier to understand.[2] In other cases it may be seen to be unfairly prejudicing the members of a jury.

1 Gregory P Joseph, *Modern Visual Evidence*; (L J Seminars Press 2009); Neal Feigenson and Christina Spiesel, *Law on Display: The Digital Transformation of Legal Persuasion and Judgment* (NYU Press 2009).
2 A M Burton, D Schofield and L M Goodwin, 'Gates of global perception: forensic graphics for evidence presentation', *Multimedia '05: Proceedings of the 13th Annual ACM International Conference on Multimedia* (ACM Press 2005) 103–11; J Mervis, 'Court views engineers as scientists' (1999) 284 (5411) Science 21.

3.92 At first glance, these computer-generated graphical reconstructions may be seen as potentially useful in any court, and they are often treated like any other form of digital evidence regarding their admissibility. In particular, they are admitted as part of expert testimonial evidence or as a special type of real evidence.[1] However, this specific form of digital media warrants special care and attention due to its inherently persuasive nature, and the undue reliance that the viewer may place on evidence presented through a (potentially photorealistic) visualization medium such as this, often to the exclusion of the underlying evidence and the assumptions made to generate these graphical representations. This is often referred to as the 'seeing is believing' tendency.[2]

1 For example, see *R v Robert Lee Clarke* [1995] 2 Cr App R 425.
2 Fred Galves, 'Where the not so wild things are: computers in the courtroom, the federal rules of evidence, and the need for institutional reform and more judicial acceptance' (2000) 13 Harv J L & Tech 161–302; Christine O Spiesel, Richard K Sherwin and Neal Feigenson, 'Law in the age of images:

the challenges of visual literacy', in A Wagner, T Summerfield and F S B Vanegas (eds), *Contemporary Issues of the Semiotics of Law* (Hart, 2005); Richard Sherwin, 'Visual literacy in action: law in the age of images', in J. Elkins (ed), *Visual Literacy in Action* (Routledge, 2007) 179–94; Damian Schofield, 'The use of computer generated imagery in legal proceedings' (2016) 9 Digital Evidence and Electronic Signature Law Review 1.

3.93 As courts begin to increasingly use multimedia and cinematic displays, this has profound implications for the legal processes taking place that are intrinsically tied to the application of such technology. It must be questioned whether the decisions made in courts when using such technology are adversely affected by this manner in which the evidence is presented.[1]

1 Ken Fowle and Damian Schofield, 'Visualising forensic data: investigation to court', in Andrew Woodward and Craig Valli (eds), *Proceedings of the 9th Australian Digital Forensics Conference* (Security Research Centre 2011); Joanna Gallant and L.auren Shepherd, 'Effective visual communication: scientific principles and research findings', in Sanuel H Solomon, Joanna Gallant and John P Esser (eds), *The Science of Courtroom Litigation: Jury Research and Analytical Principals* (ALM Publishing 2009). David M Paciocco seems to fail to have understood this serious issue when commenting that the introduction of computer enhanced photographs did not require any special evidential foundations or relevant expert evidence: 'Proof and Progress: Coping with the Law of Evidence in a Technological Age', (2013) 11 Canadian Journal of Law and Technology 181, 186–7.

Computer-generated evidence in England and Wales: civil proceedings

3.94 An early occurrence of the use of computer-generated evidence is seen in the civil case of *The Owners of the Ship Pelopidas v The Owners of the Ship TRSL Concord*.[1] In 1996 a collision took place in the Access Channel to Buenos Aires between two vessels: the *Pelopidas* and TRSL *Concord*. The issue for the court to decide was the liability for the collision and the apportionment of that liability. The items of computer-generated evidence submitted were two-dimensional computer-generated simulations of both vessels' trajectories; these were, in effect, animated maps. A 'black box' on the *Concord* recorded various positioning, speed and heading data at 15-second intervals for the relevant collision time period. Both sides accepted the accuracy of the plot. David Steel J concluded that a fair apportionment of liability was 60:40 in favour of *Pelopidas*, and stated:

> …there is a danger of losing sight of the true value of reconstructions. Of course they enable the Court and the parties to have a broad bird's eye view of the events leading up to collision. But their true probative value is that they may sometimes enable the Court to determine, not what may have happened, but what could not possibly have happened.[2]

1 [1999] 2 Lloyd's Rep 675, [1999] 2 All ER 737 (Comm).
2 [1999] 2 Lloyd's Rep 675, 682.

3.95 In stating the above, David Steel J was remarking on his accumulated experience of the usefulness of computer-generated reconstruction evidence.[1] Similar examples of the use of computer animations and simulations can also be found in *Maersk Oil UK Ltd v Dresser-Rand (UK) Ltd*[2] and *Halliburton Energy Services Inc v Smith International (North Sea) Ltd*.[3]

1 Charles Macdonald, 'Case Note *Owners of the Ship Devotion v Owners of the Ship Golden Polydinamos*' (1995) 4 Int ML 77 where the members of the Court of Appeal endorsed the comments of the trial judge respecting the use of computer simulations as evidence of a collision.
2 [2007] EWHC 752 (TCC).
3 [2006] EWCA Civ 1715.

Computer-generated evidence in England and Wales: criminal proceedings

3.96 The Court of Appeal has indicated that it favours use of digital images in criminal proceedings, as indicated by Thomas LJ in *R v Smith*:[1]

> The presentation of the evidence to the jury made no attempt to use modern methods of presentation. The presentation to this court was similar; a large amount of time was wasted because of this. It was incomprehensible to us why digital images were not provided to the jury; the refusal of NAFIS [National Automated Finger Print Identification System] to permit a digital image to be supplied to the court was a further example of the lack of a contemporary approach to the presentation of evidence. The presentation to the jury must be done in such a way that enables the jury to determine the disputed issues.[2]

1 [2011] EWCA Crim 1296.
2 [2011] EWCA Crim 1296 [61(viii)]; for New Zealand, see *R v Garrett* [2001] DCR 955 and *R v Little* [2007] NZCA 491.

3.97 However, due to the critical nature of criminal trials, it is crucial that any computer-generated evidence that is put forward be thoroughly examined.[1] The use of a jury in criminal cases is another important reason for assessing the relevance, accuracy, and possible prejudicial effect of computer-generated evidence carefully. For this reason, it is important for defence counsel to be aware of the issues that arise and be suitably prepared to test the evidence. In *R v Gardner*,[2] a person was killed during a fire in a block of flats. One of the experts who gave identification evidence for the prosecution used a new technique that deployed computer software to provide an analysis of video surveillance footage, as described by Waller LJ:

> [The expert] had developed a different technique. He had developed equipment to enable the images on a video surveillance film to be presented so as to extract as much information from it as possible. This included enhancing the film by computer to allow frame by frame examination, the ability to zoom in on part of the frame to alter the contrast and brightness to bring out detail and to run the film backwards and forwards. The second purpose of the equipment is to assist in making comparisons between one frame and another. To help in that [the expert] has developed three techniques. He called the first of them 'image addition'. By means of his computer he takes an image from one sequence of movements and selects from another sequence an image of a person who displays approximately the same stance and is about the same distance from the camera as the first. The second image is superimposed on the first so the viewer can observe whether the two images are like one another and whether there are any differences. The difference, depending on what it is, may show that the images are of different people. The second technique is referred to as 'image subtraction'. [The expert] takes the two images selected because of their comparable poses and distances from the camera and turns the first computerised image into a negative and superimposes the second on it in a positive form. The result is that the features which are common to both images disappear and only what is different remains.

> [The expert's] third technique is a 'blink comparison' whereby he can switch from one image to another. When there are differences between the two they generate an illusion of movement so that the eye is able to pick up the differences. That technique also enables the viewer to see that when one image is removed an element which had appeared to belong to the picture which has been removed in fact belongs to the picture which remains.[3]

1 For an examination of the issues and case law, see Tony Ward, 'Surveillance cameras, identification and expert evidence' (2012) 9 Digital Evidence and Electronic Signature Law Review 42.
2 [2004] EWCA Crim 1639.
3 [2004] EWCA Crim 1639, [34]; the admissibility of such evidence was approved in *R v Briddick* [2001] EWCA Crim 984.

3.98 Even though the defence did not have any material in relation to which they could cross-examine the expert witness and enable the jury to judge the expert's analysis and assessment that the person identified in the surveillance footage was the defendant, the court guardedly accepted the admissibility of this evidence. In doing so, Waller LJ also sounded a note of caution in relation to new techniques relating to identification. The judge quoted the following statement of Lord Hope in *Hopes and Lavery v HM Advocate:*

> If admitting evidence of this kind seems unfamiliar and an extension of established evidential practice, the answer must be that, as technology develops, evidential practice will need to be evolved to accommodate it. Whilst the courts must be vigilant to ensure that no unfairness results, they should not block steps which enable the jury to gain full assistance from the technology.[1]

1 [2004] EWCA Crim 1639, [45]

3.99 But even if juries are to be enabled to benefit from the full spectrum of technological evidence, they are particularly vulnerable, often more so than judges and coroners, to any prejudicial effect and inaccuracy of scientific animations. Perhaps this is because juries do not have the same level of cynicism that years of experience with analysing evidence has given judges and, to a lesser degree, coroners. In the case of *R v Ore*,[1] Tucker J stated the defence's apprehension for the admissibility of a computer-generated animation:

> The concern which is expressed by [the defence] … is as to the impact which this evidence will have upon the jury and I understand that concern. [The defence] fears that the weight which the jury may place upon the graphic animation will be disproportionate to its value in the case. [The defence] fears that they may be distracted from concentrating as they ought to do upon the evidence to be given by the expert witnesses on either side and is concerned, naturally, that the graphic animation reproduces simply one particular side of the coin.

1 (1998, unreported), (Birmingham Crown Court). Stephen Mason tried to obtain a copy of the transcript of the case for the first edition of this text, but the tapes were destroyed, in accordance with the relevant retention and disposal policy (correspondence with Michael Ives of Marten Walsh Cherer Limited). Stephen Mason subsequently corresponded with Sir Richard Tucker, who indicated that he no longer had the notes of this trial, but kindly confirmed the remarks that are attributed to him as quoted in this text.

3.100 The concerns stated above are highly relevant and illustrate real fears about any computer-generated evidence. This is especially true for forensic reconstructions. Hence, any computer-generated reconstructions should be made as precisely and in

as unbiased a way as possible, and their use has to be shown to be necessary.[1] Their probative value should outweigh any potential prejudicial effect.

1 In *Maloney v R* [2003] EWCA Crim 1373, a reconstruction was developed using computer simulation software in preparation for an appeal against conviction, a technology that was not available at the time of trial. The members of the Court of Appeal decided, in the light that the opinion of the expert that undertook the simulation was not conclusive, that the evidence would have no effect upon the safety of the conviction, and the court did not receive it and dismissed the appeal. It is not clear whether Mr Adrian Redgrave, QC, who appeared for the Crown at trial and on the reference to the Court of Appeal (Criminal Division), explored the technical integrity or the assumptions upon which the program was prepared.

3.101 These lessons may be illustrated by the case of *R v Ore*, which introduced one of the first forensic computer-generated animations to an English criminal trial. The Crash Investigation and Training Unit of the West Midlands Police Service produced the animation. The case involved a collision between two vehicles at a junction; one of the drivers was killed as he pulled out in front of an oncoming vehicle. The views of both drivers were partially obscured by large hedges and walls around the junction.[1] Tucker J, who presided over this case, further stated in his ruling on 25 November 1998:

> I am told that this is the first time in which it has been suggested that a jury in a trial such as this should be shown a computer aided animation which pictorially represents a reconstruction of a road traffic accident. It may be that in years to come such displays will be commonplace and that lawyers will marvel that anyone should ever have questioned their admissibility.
>
> … I am satisfied that it would be right to admit this evidence and, indeed, wrong to refuse so to do, provided, as I shall try to do, that I give the jury proper directions as to their approach to this evidence and provided I ensure, so far as I can, that they do not place disproportionate weight upon it. Accordingly, I rule that the evidence is admissible.[2]

1 M Doyle, 'Working model: helping the police with their enquiries' (1997) CAD User 62–63.
2 *R v Ore* (1998, unreported), (Birmingham Crown Court).

3.102 A well-known example from Northern Ireland is the computer-generated evidence that was extensively used during the Bloody Sunday Inquiry.[1] In 1972, 13 people were killed during a peaceful demonstration. The original inquiry produced a report within 11 weeks of the incident, and acquitted the soldiers involved. In 1998, a Tribunal of Inquiry was established to reassess the events.[2] Lord Saville, the chair of the tribunal, took full advantage of ensuing improvements in technology, and used a computer software system designed especially for use in the Inquiry to amplify the testimony of witnesses. The Northern Ireland Centre for Learning and Resources produced the computer-generated virtual models, which reconstructed a large area of Londonderry that had been extensively altered since 1972. The user was able to compare the same scene as it appeared at the time of the Inquiry and as it was in 1972. There were 80 locations stored in the system that could be explored, with specific points of view being recalled when switching between the representations. The system could also store oral evidence about location and movement, and export scenes to a mark-up system so that witnesses could draw on top of images. The computer software system that was admitted was deemed to be unbiased and accurate.

1 See 'The Bloody Sunday Inquiry' <www.bloody-sunday-inquiry.org.uk>.

2 Statement by Tony Blair, Prime Minister: HC Deb 29 January 1998, vol 305, col 501.

3.103 The Bloody Sunday Inquiry computer system was not interactive in three-dimensions. Virtual reality or VR, by definition, is an interactive computer-generated simulated environment with which users can interact using a computer monitor or specialized hardware. The computer system used for the Bloody Sunday Inquiry was interactive in the sense that viewers were able to view images of different scenes at varied times. However, the viewer was not able to move around a full three-dimensional virtual environment of Londonderry itself, since the full three-dimensional virtual model of the area did not exist. But over the last few years courts in England and Wales have begun to introduce interactive three-dimensional VR crime scene environments for a number of high profile criminal cases.[1] There is little doubt that with the increasing complexity of criminal investigations, we will see more use of virtual environments and immersive virtual environments in legal proceedings.

1 Damian Schofield, 'Playing with evidence: using video games in the courtroom' (2011) 2 Journal of Entertainment Computing 47.

3.104 Virtual environments possess the potential to sway juries and decision makers, even more so than computer animations in general. Creating an environment thatt allows viewers to take different perspectives and manipulate objects in that environment do indeed allow for 'what-if' scenarios to be played out, and could lead to more robust decisions. But the reconstructions of scenes in these environments are based on various assumptions and premises, not all of which can be elucidated or are transparent, or easily accessible for review by opposing experts and by decision makers. Indeed, analyses of computer-generated displays show that they can be extremely advantageous in the court, provided they are used appropriately. The consequences of a failure to investigate these issues cannot be underestimated, since errors, inaccuracies, misuse, tampering or biases within visualizations are capable of leading to miscarriages of justice.[1]

1 Marcel Worring and Rita Cucchiara, 'Multimedia in forensics', in *Proceedings of the 17th ACM International Conference on Multimedia* (ACM Press 2009) 1153–1154.

4

Hearsay

Chris Gallavin and Stephen Mason

4.1 The much maligned evidential rule of hearsay exclusion has been subject to some interesting challenges in many common law jurisdictions over the past 15 years. An anathema to lawyers of the civil or administrative law system and seemingly largely misunderstood in its complexity by many common law lawyers, the hearsay rule has been so undermined as to bring into question its continued existence. This chapter does not provide a comprehensive exposé of the hearsay rule. However, in drawing the rule back to his historical foundation we will, in part, question its relevance in the context of digital evidence and attempt to universalise the considerations that are at play in the admission of second-hand evidence of a digital nature.

4.2 The hearsay rule of exclusion is a rule that has long been considered a complex and confusing exclusionary rule of evidence.[1] Whilst seemingly a central tenet and peculiarity of the adversarial system of justice, we suggest the underlying premises of testability, reliability and weight remain universal for both the common law and the civil law systems. Considering the use of the word 'testability', Alex Stein considered this when propounding the 'principle of maximal inferential individualization'. Stein suggested:

> (1) No adverse inference should be drawn against the defendant, unless it has been exposed to and survived the maximal individualized testing;
>
> (2) This includes every practical possibility of testing the applicability of the inference in question to the individual defendant's case;
>
> (3) The defendant should accordingly be provided with appropriate immunities from the risk of error.
>
> …
>
> When two inquiries may be directed to the same end, evidence commencing the more promising inquiry should preempt the evidence activating the less promising alternative. Judges should therefore follow the 'best evidence principle', which would exclude secondary evidence when better evidence is available. 'Better evidence' would be that which enables judges to reach its probandum in a fewer inferential steps. By saying this, I refer not merely to the degree of the logical directness of the evidence vis-à-vis its probandum, but also, and, indeed, primarily, to the extent of its testability. Evidence giving rise to transforming arguments that can be examined, and thus strengthened or weakened, with greater ease should always be preferred. This principle would ascribe preferability not merely to original evidence, as opposed to its duplicate… [footnotes omitted][2]

1 Colin Tapper, 'Reform of the law of evidence in relation to the output from computers' (1995) 3 Intl J L & Info Tech 79 for a critique and suggestion that the rule should be abolished. In 1989, the New Zealand Law Commission summarised that the rule of hearsay exclusion and its exceptions were, 'unclear, inconsistent, and lacking in coherence' (Law Commission, *Hearsay Evidence: An Options Paper* (NZLC PP10, 1989), p. vi).

Chris Gallavin and Daniel Seng, 'Hearsay', in Stephen Mason and Daniel Seng (eds.), *Electronic Evidence* (4th edn, University of London 2017) 70–87.

2 Alex Stein, 'The refoundation of evidence law' (1996) 9 Journal of Law and Jurisprudence 279, 326–7, 331.

4.3 In the context of electronic evidence, this must be right, and both George L. Paul and Steve W. Tepper have argued that 'testability' of a digital system is now essential.[1]

1 George L Paul, 'Systems of evidence in the age of complexity' (2014) 12 Ave Maria Law Review 173; Steve W Tepper, 'Testable reliability: a modernized approach to ESI admissibility' (2014) 12 Ave Maria Law Review 213.

4.4 It is these fundamental issues, together with an ill-defined and oft misunderstood 'right to confront', that have stood as the historical foundation for a rule that has lost some of its force at best and is out-dated at worst. Aside from this public policy consideration, which we consider to have little relevance in all non-United States common law jurisdictions, the continued existence of hearsay as a general rule of exclusion falls to truth-finding factors for its survival. These considerations include the fact-based issues of authenticity, reliability, relevance and weight. If these considerations can be addressed either by their being substantively satisfied or through establishing a means of testability that may lead to their satisfaction, then exclusion of hearsay is not warranted.[1]

1 As the question of admission only requires the consideration of a threshold level of reliability, a court need not concern itself with establishing whether in fact the evidence is reliable.

4.5 The complexity of the hearsay rule is increased in the case of digital evidence. First, a distinction needs to be made between statements capable of being hearsay and evidence not meeting the definition of a hearsay statement, the latter resulting in the evidence being treated not as hearsay but as real evidence. While a communication written by a person and stored in a digital form is capable of being a statement for the purpose of the hearsay rule, 'statements' derived from software code are not. In the case of the latter, where raw data is entered into a program and then processed by digital means, the resulting apparent statement may not qualify as a statement for the purpose of the hearsay rule. A distinction therefore needs to be drawn between the content of text messages and emails, the presentation of raw data (for example, the metadata in Gmail, Hotmail, graphs, charts, presentations etcetera), and the presentation of information derived as a consequence of action by software code (for example, conclusions resulting from predictive logic and the presentation of computer generated conclusions and advice drawn from data). Second, the issues of authenticity and the application of the traditional exception relating to business records will also form two important touchstones in the application of the rule.[1] Third, reliability, whilst a consideration in the context of relevance (a fact needs a semblance of reliability to be relevant),[2] is also an issue in the application of what is now the main exception to the hearsay rule across all common law jurisdictions – apparent reliability,[3] especially in relation to automatically produced records as circumstantial evidence.[4] The particular application of these principles to electronic evidence illustrates that for the hearsay rule, the treatment of electronic evidence is complex and will raise unique issues that will often make electronic evidence stand apart from other forms of evidence.

1 Authenticity and the business document exception are two items dealt with elsewhere in this book, for which see the chapter on authentication.
2 See *R v Bain* [2009] NZSC 16.

3 In New Zealand this is referred to as a reasonable assurance of reliability (Evidence Act 2006, s 18). In Canada, the test is referred to as a 'circumstantial guarantee of trustworthiness'; see *R v Starr* 2000 SCC 40, see also *R v Khelawon* 2006 SCC 57. See also Evidence Act 1995 (Cth), s 65.
4 *R v Davis* [2006] EWCA Crim 1155, [2007] Crim LR 70 (note), use of a mobile telephone; *R v Bailey* [2008] EWCA Crim 817, evidence of a chatroom.

The foundations of the rule of hearsay exclusion

4.6 It is interesting to begin with a traditional and simple definition of the hearsay rule. Sir Rupert Cross defined the hearsay rule of evidence as '[A] statement other than one made by a person while giving oral evidence in the proceedings is inadmissible *as evidence of any fact stated*.'[1] In offering this definition, Sir Rupert Cross combined the notion of a particular form of statement with the necessity to exclude.[2] Although simple in its expression, this definition has proved unhelpful in its application. This is because the historic exceptions are so numerous as to warrant the general principle near void. Importantly, it does nothing to define 'statement', and in light of the modern move away from including implied assertions within the hearsay rule,[3] a contemporary definition of the hearsay principle would probably be somewhat different.

1 Rupert Cross, *Evidence* (5th edn, Butterworths 1979) 6. In his first edition, Phipson stated that hearsay was 'Oral or written statements made by persons not called as witnesses are not receivable to prove the truth of the matters stated' (Sidney L Phipson, *The Law of Evidence* (Stevens and Hayes 1892) 117). See also the definition suggested by Charles Cato who preferred to see hearsay limited to 'unsworn utterances containing narrative assertion, where it is a suggestion for reform' ('Verbal acts, res gestae and hearsay: a suggestion for reform' (1993) 5 Bond Law Review 72, 73).
2 See below for a discussion of hearsay statement.
3 See Australian Law Reform Commission, *Uniform Evidence Law* (Report No 102, 2006) paras 7.19–7.22, <www.alrc.gov.au/publications/7.%20The%20Hearsay%20Rule%20and%20Section%2060/unintended-assertions>.

4.7 The rationale for the exclusion of hearsay evidence has been put succinctly by Allan, who states that '[t]he basic rationale of the hearsay rule rests on the right of cross-examination'.[1] Without the benefit of cross-examination, there exists a perception that evidence will be subject to at least four clear risks:[2] the weakness of any witness perception, the weaknesses in recording and later recollecting that perception, the problem of narration or the portrayal of the recollected perception, and the risk of a lack of witness sincerity and the possibility of fabrication.[3] Each of these represents *risks* rather than the guaranteed demonstration of problems undermining the reliability of evidence. The significant issue is that in the absence of cross-examination, the common law is reluctant to rely on the accuracy and therefore reliability of second-hand evidence[4] in the task of assigning weight to evidence.[5]

1 James Allan, 'The working rationale of the hearsay rule and the implications of modern psychological knowledge' (1991) 44 Current Legal Problems 217. On the dangers of hearsay evidence see E M Morgan, 'Hearsay dangers and the application of the hearsay concept' (1948) 62 Harvard Law Review 177, 178–9. On the perceived virtues of cross-examination, see 2 Bl Comm 373, where Sir William Blackstone stated that examination through '*viva voce*, in the presence of all mankind, is much more conducive to the clearing up of truth', and Matthew Hale, *History and Analysis of the Common Law of England* (J Nutt, 1713) 258 <www.constitution.org/cmt/hale/history_common_law.htm> (cited in the US Supreme Court case of *Crawford v Washington* 541 U.S. 36 (2004)), where it is said that cross-examination 'which beats and boults out the Truth much better than when the Witness only delivers a formal Series of his Knowledge without being interrogated'. The Supreme Court of Canada proclaimed cross-examination as 'the optimal way of testing testimonial evidence', *R v Khelawon* 2006 SCC 57 [35].

2 See John H Wigmore, *A Treatise on the Anglo-American system of Evidence in Trials at Common Law* (3rd edn, Little, Brown 1940) para 478. See also Edmund Morgan, 'Hearsay dangers and the application of the hearsay concept' (1948) 62 (2) Harv L Rev 177; Laurence H Tribe, 'Triangulating hearsay' (1974) 87 *Harv L Rev* 957; and Michael H Graham, 'Stickperson hearsay: a simplified approach to understanding the rule against hearsay' (1982) 4 *University of Illinois Law Review* 887.
3 Edward W Cleary (ed.), *McCormick on Evidence* (West Publishing 1984) para 245.
4 That is, evidence that is more than one remove from the first statement, or 'irrespective of the number of intermediate communications between the original source and the testifying witness': Colin Tapper, *Cross & Tapper on Evidence* (12th edn, Oxford University Press 2010) 552 fn 9; the authors of Australian Law Reform Commission, *Uniform Evidence Law* (Report No 102, 2006) refer throughout to 'second-hand' hearsay evidence.
5 Some however, have significantly undermined the importance of cross-examination; see in particular Elisabeth McDonald, 'Going "straight to basics": the role of Lord Cooke in reforming the rule against hearsay – from Baker to the Evidence Act 2006', (2008) 39 VUWLR 143.

4.8 Acknowledging the unworkability of the traditional definition of hearsay noted above, the hearsay rule has been amended in two jurisdictions in particular, England & Wales and New Zealand.[1] The universal theme of reform has been the diminution of the influence of the rule with the object of allowing for increased admission of hearsay evidence. In reforming the application of the rule, these jurisdictions have reassessed the foundation of the hearsay rule by refocusing the rule and rationalising the myriad exceptions that had increased to the point of confusion.

1 See the New Zealand Evidence Act 2006.

4.9 One of the central pillars of the hearsay rule is the principle that evidence that cannot be tested through application of the traditional mechanisms of the adversarial process is not to be trusted. The traditional approach was therefore to exclude suspect evidence, not because a court could not be assured of its reliability, but that its reliability was effectively unknowable due to the absence of an ability to cross-examine. Hence the term 'second-hand evidence'. The influence of reliability as an exception to the rule, together with a more expansive approach into the ways in which such reliability can be established, has focused the common law to look for indicia of apparent reliability and expanded its willingness to consider alternatives to cross-examination conducted in the context of a proceeding.

Public policy justifications for a rule of exclusion

4.10 A further, apparently important but ill-defined foundation of the hearsay rule is the public policy consideration of the right to confront an accuser. Although expressly established in the Constitution of the United States, there is no 'right' *per se* of confrontation in other common law jurisdictions. That is not to say that the notion has no influence, but that its authority is more amorphous and indefinable in the context of the regular rule of hearsay operating across the common law world. In the digital context, this historic foundation emphasizes the human-centric nature of statements. It is a misnomer that a statement that is wholly electronically derived ought not to be a statement capable of supporting the application of the hearsay rule in relation to the notion of confrontation.

4.11 As a foundation for the rule against hearsay, the right to confront draws on the notion that the right to humane treatment and procedural integrity both *feel*

undermined by the admission of hearsay evidence.[1] However, we question the supposed legitimacy that this right gives to the continued existence of a hearsay rule of exclusion. There is no justification for the exclusion of evidence that can otherwise be assured to be reliable because it is not presented in a form that allows an accused to confront his accuser.

1 For discussion of the foundation of this right and its modern legitimacy, see Mike Redmayne, 'Confronting confrontation', in Paul Roberts and Jill B Hunter (eds), *Criminal Evidence and Human Rights: Reimagining Common Law Procedural Traditions* (Hart 2012) 283.

4.12 In addition, this public policy justification for the rule of exclusion is inherently amorphous, difficult to define and equally difficult to assign importance to. As highlighted by Redmayne, this justification may be based upon notions as varied as 'the accuser has an obligation to face the accused' and 'it is not the way things are done'.[1] In light of such intangibility, it is difficult to determine whether the rule of law, for example, demands the giving of testimony in open court in some or all cases – notwithstanding the fact that sometimes the out-of-court statement of a witness that is not available to give evidence is manifestly reliable or capable of adequate testing independently of cross-examination. In this respect, a divergence in approach has arisen between on the one hand, the United States and the European Court of Human Rights, and on the other, the rest of the common law world.

1 Redmayne, 'Confronting confrontation' 296. See also Toni M Massaro, 'The dignity value of face-to-face confrontations' (1998) 40 University of Florida Law Review 863.

4.13 Policy considerations in the context of the right of confrontation take their fullest form under the United States Constitution. Ratified on 15 December 1791, the Sixth Amendment provides a criminal defendant with the right to a speedy trial and the ability to confront witnesses. The right 'to be confronted with the witnesses against him' was included in response to the trial of Sir Walter Raleigh in 1603. A brief examination of this trial reveals valuable insights into the meaning and intention of the right to confront as included in the United States Constitution.

4.14 On 17 November 1603, Sir Walter Raleigh was tried for high treason for his part in the 'Main' or 'Spanish Treason' conspiracy to murder King James I.[1] The conspiracy was to place Lady Arabella Stuart on the throne and was said to involve Raleigh, Cobham and George Brooke. George Brooke, the brother of Lord Cobham, has been described as 'a man sensible and well educated, but turbulent and totally unprincipled'.[2] Similarly, Cobham has been described as 'a man of extremely weak intellect', and that at his own trial, 'he exhibited the most contemptible baseness and cowardice'.[3] It would appear that the trial of Raleigh was politically motivated, with Robert Lord Cecil being said to have pushed heavily for his prosecution.[4] At trial, statements made by Lord Cobham to the Privy Council and in a letter were adduced in evidence of the existence of a conspiracy in which Raleigh was leader. Lord Cobham was not called to give evidence before the jury in person. In answer to the failure to call Lord Cobham, Raleigh stated:

> But it is strange to see how you press me still with my Lord Cobham, and yet will not produce him; it is not for gaining of time or prolonging my life that urge this; he is in the house hard by and may soon be brought hither; let him be produced, and if he will yet accuse me or avow this Confession of his, it shall convict me and ease you of further proof.[5]

1 At trial, the basis upon which an intention to kill the King was the evidence of an overheard conversation in which a nobleman had said when referring to Lord Cobham that 'there was no way of redress save by taking away the King and all his cubs': David Jardine, *Criminal Trials*, I (Charles Knight 1832) 395.
2 Jardine, *Criminal Trials* 390. Both Lord Cobham and Brooke were implicated in a previous unsuccessful Catholic plot to kidnap the King and extract certain proclamations from him including tolerance of the Catholic faith. However, as Jardine observed, the plot was '… so absurd, and composed of so many elements of discord, and to be executed by persons who … agreed in nothing but their common discontent, contained within itself the seeds of dissolution' (392).
3 Jardine, *Criminal Trials* 394.
4 Jardine, *Criminal Trials* 394.
5 Jardine, *Criminal Trials* 427.

4.15 In highlighting the trial of Sir Walter Raleigh, the United States Supreme Court in *Crawford v Washington*[1] noted that it was this type of case and accusation that the right of confrontation was intended to serve. Although many limitations have arisen in the application of the United States' right of confrontation, the principle remains that we may feel uncomfortable admitting second-hand evidence – not only because of its possible unreliability, but because it seems to be opposite the ethos of a fair fight. This may remain as a belief that encourages us to present the best evidence wherever possible, but in cases where there exists proof of reliability, the notion of a right to confront ought to be reconsidered in jurisdictions where such a formal 'right' is not recognized.

1 541 U.S. 36 (2004).

4.16 The notion of a right to confront has consequences for the way in which courts deal with evidence in digital form. First, this historical ground of justification illustrates that an offending hearsay statement is one which comes from a human source. If the statement is the product of automation, software calculation or predictive logic, then using this ground of justification, it cannot be said to be 'a statement' for the purposes of the hearsay rule. There is no opportunity for a human agent to provide the statement to the court because no single human was responsible for the communication. Second, the nebulous notion of a right to confrontation and the inability to apply it in some form may underline the hesitation of some judges in dealing with electronic evidence, particularly where the offending statement meets the traditional criteria of hearsay.

Defining hearsay

4.17 Significant erosion of the hearsay rule under the common law has centred upon the identification of the scope of a hearsay statement. No longer do implied or unintended assertions fall under the definition of statement for the purpose of the hearsay rule.[1] In turning from the position first established in *Wright v Doe d Tatham*,[2] the law in England & Wales and later New Zealand and other common law jurisdictions has significantly limited the scope of the hearsay rule by restricting the definition of 'statement' to express assertions and conduct within which an intention to assert could be established.[3] Therefore, a statement is only one to which the hearsay rule can apply if, and only if, an intention to communicate can be identified within that statement or conduct.

1 For the virtues or otherwise of this position, see Brenda Marshall, 'Admissibility of implied assertions: towards a reliability-based exception to the hearsay rule' (1997) 23 Monash University Law Review 200.
2 (1837) 7 A & E 313, 11 ER 1378. This position was later affirmed in the case of *DPP v Kearley* [1992] 2 AC 228 (HL). For the Australian context, see the comments of McHugh J in the Australian case of *Pollitt v R* (1992) 174 CLR 558 at [21]. Similarly, in New Zealand, see *R v Mokaraka* [2002] 1 NZLR 793 (CA).
3 See for example the definition of 'statement' in New Zealand: Evidence Act 2006, s 4; in Australia, Evidence Act 1995 (Cth), s 59(1), and England & Wales, Criminal Justice Act 2003, s 115(3).

4.18 In the context of electronic evidence, lists of figures or disparate facts may not be accompanied by an intention to communicate and therefore will not qualify as statements as defined for the purposes of the exclusionary rule. Express statements contained within a mobile telephone via text, or a computer via email, will often qualify as statements – their form of creation, capture and storage being the only difference with paper based documentation.[1] Limiting a qualifying hearsay statement to express assertions or where there is a clear intention to assert will significantly limit the application of the hearsay rule, thereby placing much digital evidence beyond the realm of the exclusionary rule.

1 Aside from issues of authenticity which will often present particularly unique considerations. See chapter 7 on authentication.

Civil proceedings and the requirement to give notice

4.19 The hearsay rule under the common law has significantly receded over the past 15 years, in large part because of the narrow definition of a hearsay statement. The hearsay rule provides that only a witness giving evidence could testify to the truth of the assertions he made in evidence. In England & Wales, the hearsay rule was abolished for civil proceedings by s 1(1) of the Civil Evidence Act 1995. The Act applies to all civil proceedings,[1] including proceedings in the magistrates' court.[2] By contrast, New Zealand relies upon an expansive reliability exception to the rule rather than differentiating between criminal and civil jurisdictions.[3]

1 Civil Evidence Act 1995, s 11.
2 The Magistrates' Courts (Hearsay Evidence in Civil Proceedings) Rules 1999, SI 1999/681.
3 In terms of reliability as applied under the Canada jurisdiction, see the Canadian Supreme Court in *R v Khelawon* 2006 SCC 57, [2] Cf. *Horncastle v R* [2009] EWCA 964, [57].

4.20 In England & Wales, a party that intends to adduce hearsay evidence in civil proceedings is required to give the other party or parties notice of his intention and, should it be requested, particulars of the evidence.[1] This requirement to give notice is not unique to England & Wales. A criticism of hearsay evidence said to justify the existence of a rule of exclusion is that admission of hearsay would amount to an unjustified element of surprise causing delay and unwarranted disruption in a proceeding.[2] This criticism has largely been addressed through the need to give notice of an intention to call a witness.[3] In New Zealand, where the requirement for notice can be waived by a judge, it has been suggested in one case that the nature of the statement as hearsay evidence could in effect be overlooked by a determiner of fact by considering the reliability that comes with a 'course of business'.[4] Although the case involved the purchase of a precursor chemical to the manufacture of methamphetamine, the Court

gave an analogy of purchasing petrol at a service station. According to the members of the New Zealand Court of Appeal, a customer purchasing petrol was not relying on the label on the bowser (fuel dispenser), but rather the course of business giving surety to the fact that it was petrol and not diesel coming from the bowser. If this analogy was relied upon in the context of digital evidence, an entirely new rule would develop. It would not be as an exception to the hearsay rule, but in parallel to the rule, causing difficulty not only in differentiating the circumstances in which the hearsay rule would or would not apply, but potentially undermining authenticity and the business document exception. A more robust approach is to acknowledge the hearsay value of the document or electronic record, and then directly address the issue of notice, and if notice is not given, consider whether a waiver is in the interests of justice, rather than attempting to devise a further exception to or subversion of the hearsay rule.

1 Civil Evidence Act 1995, s 2.
2 See Chris Gallavin, *Evidence* (LexisNexis 2008) 127. The irony of this justification for the rule of exclusion is that argument over the application of the rule was likely to lead to more delay and greater expense as would otherwise have been the case.
3 In the context of New Zealand, see Evidence Act 2006, s 22; in Australia, Evidence Act 1995 (Cth), s 67.
4 *R v Lenaghan* [2008] NZCA 123. See also Chris Gallavin, '*R v Lenaghan:* is it business as usual in New Zealand despite the reforms of the Evidence Act 2006?' (2008) 12 E & P 325.

4.21 Returning to the Civil Evidence Act 1995 (as an example of an approach to hearsay applicable to digital evidence across the common law), the Act includes a number of exceptions to the hearsay rule that are particularly relevant to documents stored in digital form. Published works dealing with matters of a public nature, public documents and public records are all admissible under the provisions of s 7(2) of the Civil Evidence Act 1995 – similar to provisions operating in other jurisdictions. More distinctively, where a document can be shown to be part of the records of a business or public authority, the document can be received into evidence in civil proceedings without further proof in accordance with s 9. The wording of this and similar provisions in other jurisdictions means that the form a technology takes will not prevent the admission into evidence of data stored in digital form.

Criminal proceedings

4.22 The right of confrontation under the United States Constitution only applies to criminal proceedings, and evidence that is testimonial in nature. For all other evidence, the hearsay rule applies, with its board application of the reliability exception. As stated above, in England & Wales, the enactment of the Criminal Justice Act 2003 repealed the provisions relating to hearsay in the Criminal Justice Act 1988, and by doing so, reversed the decision on implied assertions in the case of *DPP v Kearley*,[1] as well as abrogated most of the common law of hearsay.[2] The operative provision is s 114(1), which reads:

> Admissibility of hearsay evidence
>
> (1) In criminal proceedings a statement not made in oral evidence in the proceedings is admissible as evidence of any matter stated if, but only if—
>
> (a) any provision of this Chapter or any other statutory provision makes it admissible,

(b) any rule of law preserved by section 118 makes it admissible,

(c) all parties to the proceedings agree to it being admissible, or

(d) the court is satisfied that it is in the interests of justice for it to be admissible.

1 *DPP v Kearley* [1992] 2 AC 228 (HL).

2 Previously, where a computer recorded the numbers of various components that were fitted to motor cars, the print-out was a hearsay statement where it was offered in evidence to prove that a number of components were fitted to a specific motor car: *Myers (James William) v DPP* [1965] AC 1001 (HL); Michael Hirst, 'Hearsay, confessions and mobile telephones' (2011) 75 Journal of Criminal Law 482, 483.

4.23 The provisions of s 114 serve as an introductory provision to the other provisions in that chapter. Section 114 retains the exclusion of the hearsay rule,[1] but operates to admit hearsay statements in criminal proceedings within the parameters set out in (a) – (d) (although a number of common law exceptions are retained by virtue of s 118). In addition, s 121 provides for additional requirements for the admissibility of multiple hearsay,[2] and s 126 provides for the general discretion to exclude evidence.

1 Tapper, *Cross & Tapper on Evidence* 602.

2 Evidence from a Police Incident Log was wrongly admitted under s 117 at trial, but on appeal, the members of the court decided that the evidence was correctly admitted under s 121(c), in *Maher v DPP* [2006] EWHC 1271 (Admin).

4.24 Of particular relevance to electronic evidence is s 129. It reads:

> 129 Representations other than by a person
>
> (1) Where a representation of any fact—
>
> (a) is made otherwise than by a person, but
>
> (b) depends for its accuracy on information supplied (directly or indirectly) by a person,
>
> the representation is not admissible in criminal proceedings as evidence of the fact unless it is proved that the information was accurate.
>
> (2) Subsection (1) does not affect the operation of the presumption that a mechanical device has been properly set or calibrated.'

4.25 The UK Law Commission considered the admissibility of a computer print-out, whether it is hearsay, and whether the print-out itself is relevant:

> The question is, on what basis should such evidence be excluded? One view is that it is hearsay, because it is tantamount to a statement made by the person who fed the data into the machine. An alternative view is that the statement by the machine, properly understood, is conditional on the accuracy of the data on which it is based; and that, if those data are not proved to have been accurate, the statement therefore has no probative value at all. The question of hearsay does not arise, because the statement is simply irrelevant.
>
> We believe that the latter view is closer to the truth, and that it is therefore unnecessary to complicate our hearsay rule by extending it to statements made by machines on the basis of human input. On the other hand we do not think it would be safe to assume that everyone will share this view. We must anticipate the argument that, if such statements are inadmissible at present, that is because they are hearsay; that, under our recommendations, they would no longer be hearsay, because our formulation of the rule would apply only to representations made by people; and that they would therefore cease to be inadmissible.[1]

1 Law Commission, *Evidence in Criminal Proceedings: Hearsay and Related Topics* (Law Com No 245, 1997) paras 7.48–7.49.

4.26 Notwithstanding the other routes of admissibility in s 114(1), one particularly wide[1] route is to admit hearsay evidence 'in the interests of justice' under s 114(1)(d),[2] subject only to the conditions in s 114(2). However, a number of cases that deal with the inclusion of evidence of telephone calls and text messages sent on mobile telephones, especially in relation to cases involving illegal drugs, have caused some confusion. For instance, in *R v Chrysostomou*,[3] the trial judge admitted four text messages apparently sent to the appellant by someone called 'John' who attempted to set up a supply of drugs to provide evidence that the appellant was a dealer in drugs. In giving judgment for the court, Aikens LJ agreed that the text messages were not caught by the statutory code on hearsay on the basis that the messages were adduced, not to prove, as fact, any matters stated in the messages, but 'as evidence of an underlying state of affairs, which was the basis on which "John" apparently sent the texts to the appellant, namely that the appellant dealt with drugs and so could meet John's demands.'[4] In his commentary, Professor Ormerod agreed with the conclusion reached by Aikens LJ but disagreed with the reasoning, pointing out that the text messages were actually relied upon for the truth of the implied assertion contained in the message that the accused was a dealer in illegal drugs. This, however, did not render the message hearsay because, as Professor Ormerod noted that for a statement to be hearsay, the purpose of making the statement must be to cause another to believe the matter or to act on the matter a stated,[5] but, 'the purpose of the texter ['John'] was *not* to cause [the appellant] C to believe/act on his being a dealer.'[6] (emphasis added) This must be right.[7] Additionally, Professor Hirst observed that if there is nothing to prove an established relationship, or an incriminating response or reaction from the defendant, it may be inadmissible, regardless of whether it is hearsay or not.[8]

1 See e.g. *R v Humphris* [2005] EWCA Crim 2030, [11].
2 *R v Xhabri* [2005] EWCA Crim 3135; however, note the commentary (and references to other relevant articles) by Tom Worthen, 'The hearsay provisions of the Criminal Justice Act 2003: so far, not so good?' [2008] Crim LR 431; Roderick Munday, 'Athwal and all that: previous statements, narrative, and the taxonomy of hearsay' (2010) 74 Journal of Criminal Law 415; Michael Stockdale and Emma Piasecki, 'The safety-valve: discretion to admit hearsay evidence in criminal proceedings' (2012) 76 Journal of Criminal Law 314.
3 [2010] EWCA Crim 1403, [2010] Crim LR 942 (note).
4 [2010] EWCA Crim 1403, [28].
5 Criminal Justice Act 2003, s 114(1) read with s 115(3)(a), (b).
6 [2010] Crim LR 942 (note), 944.
7 See the analysis of this precise point by Professor Ormerod at [2010] Crim LR 938–941, in which he cites *R v Singh* [2006] EWCA Crim 660, [2006] Crim LR 647 (note); *R v Mayers* [2008] EWCA Crim 2989; *R v Leonard (Mark Alan)* [2009] EWCA Crim 1251, [2009] Crim LR 802 (note); *R v Fox* [2010] EWCA Crim 1280; *R v Bains* [2010] EWCA Crim 873, [2010] Crim LR 937 (note); regarding inferences to be drawn from the absence of an entry on a record, see *R v Shone* (1983) 76 Crim LR 72; M Khan, 'Hearsay' (1984) 48 Journal of Criminal Law 25–27.
8 Hirst, 'Hearsay, confessions and mobile telephones' 491 fn 25, citing *R v William O'Connell* [2003] EWCA Crim 502.

4.27 With the abolition of implied or unintended assertions from the scope of the hearsay rule, not any assertion made with the intention to communicate will be a qualifying hearsay statement. The inadmissible hearsay assertion has to be associated with the object for which it is tendered in evidence in support, failing which it is

admissible as an implied or unintended assertion. This illustrates the fundamental weakness of the rule. By excluding unintended assertions, there arises a possibility that arbitrary limits may arise in that the difference between a hearsay statement and a non-hearsay statement will rest with the question of whether there exists an intention to communicate. The existence of an intention to communicate is of such little value as to render the distinction meaningless. Furthermore, such a distinction exposes the application of the exclusionary rule to the formulation of a clever submission of a lawyer in that the application of the rule might be avoided by classifying the statement as a reflection of the mindset of the maker as opposed to an intention of the maker. In such a case, no real distinguishing factor truly exists.

4.28 Careful consideration needs to be made of the provisions of s 114(2) regarding evidence in digital form when it is obtained from the Internet and where the evidence relating to the material, such as its authorship and ownership of the web site from which it originates, is not known, as in the case of *Bucknor v R*.[1] In this case, the trial judge admitted evidence found by the police on a BEBO page, consisting of 46 separate 'pages', on the website www.bebo.com. The material included a number of photographs of Bucknor that he had taken of himself after he had left prison. The photographs had been placed on the page by someone in such a manner as to portray Bucknor as a member of the Organised Criminals (OC) gang. There was a hyperlink to a YouTube page that portrayed the OC gang as violent. The YouTube page, which was recorded on a DVD, was also shown to the jury. The prosecution did not have any evidence of the IP address from which the material was uploaded. The trial judge admitted the evidence as part of the background to the case, but on appeal, the appellant argued that the judge failed to give a sufficient direction regarding the ownership of the web site in question. The members of the Court of Appeal agreed with the submission. The material was clearly hearsay because it seemed likely that the maker as the source of the material was representing as fact or opinion that Bucknor was a member of the OC gang. In considering the issues set out in s 114(2), Hooper LJ, giving the judgment for the Court, said, that the judge ought to have considered how reliable the maker of the statement was (sub-paragraph (e)), whom the judge failed to identify.[2] Failing to identify the maker meant that it was not obvious how many levels of hearsay were involved. The judge also failed to consider the reliability of the statement that the appellant was a member of the OC. Hooper LJ concluded:

> 44. Furthermore it seems to us on the facts of this case that the judge should have considered how reliable the statement was. He should also have asked whether the prosecution could call the maker of the statement and if not why not.
>
> 45. In our view the judge did not approach section 114 as he should have done. In any event, as we have said, his direction to the jury invited them to reach conclusions which no reasonable jury could have reached.[3]

1 [2010] EWCA Crim 1152.
2 [2010] EWCA Crim 1152, [42]–[43].
3 [2010] EWCA Crim 1152, [44]–[45].

Elements of hearsay

4.29 Professor Pattenden suggests that 'A statement may be probative of a disputed fact not because of what it states (expressly or by implication) but because of what it is possible to infer from the fact that it was said (or written)'.[1] While this is undoubtedly

true, the difficulty lies in the translation of this rule into the realm of hearsay. A statement does not become hearsay nature merely because it may have probative value. As stated earlier, a statement will be capable of attracting the hearsay rule only if it encapsulates an intention to communicate and is adduced for the same purpose or object as the communication. It is not enough that it merely communicates something, or anything; this does not render it a hearsay statement as a matter of course.

1 Rosemary Pattenden, 'The rule against hearsay', in Hodge M Malek (ed.), *Phipson on Evidence* (18th edn, Sweet & Maxwell 2013) paras 28–32.

4.30 In his commentary to *R v Leonard (Mark Alan)*,[1] Professor Ormerod described four elements that establish that a statement is hearsay, as constituted by ss 114 and 115 of the Criminal Justice Act 2003:

> 1. a statement (i.e. a representation of fact or opinion) made by a person (not made automatically by a machine, if so, see s.129).[2]
>
> 2. made otherwise than in the course of the present proceedings (even testimony in previous proceedings is caught);
>
> 3. relied on by the party seeking to adduce it at trial to prove the "matter stated" and not simply that the statement was made or for some other purpose;
>
> 4. where the purpose (or one of the purposes) of the maker must have been to cause someone to believe the 'matter stated' (i.e. that content of the statement now relied on at trial) or to act upon that matter stated.[3]

1 [2009] EWCA Crim 1251, [2009] Crim LR 802 (note).
2 Although not expressly provided for under the Evidence Act 2006, the application of the hearsay rule in New Zealand undoubtedly relies on the same premise – that the statement be the result of conscious human thought.
3 '*R v Leonard (Mark Alan)*' [2009] Crim LR 802 (note), 804.

4.31 In *R v Leonard (Mark Alan)*,[1] the members of the Court of Appeal (Criminal Division) determined that two text messages sent by unknown people to the appellant on two separate mobile telephones were hearsay evidence, and should not have been admitted at trial. The content of the messages are set out as follows:

> The first, timed at 10.24 on 2nd May 2008, reads:
>
> 'Cheers for yday! Well sound gear:-S! feel well wankered today!'
>
> The second text message was from a different phone number and was on the second mobile phone. It was timed at 10.51 on 6th May 2008. It read:
>
> 'Mark, that was a proper dog cunt move mate, that joey was a £5 joey and that was my last £10. Thanks. I dont why I think u would not do that 2 me. I dont.'[2]

1 [2009] EWCA Crim 1251, [2009] Crim LR 802 (note).
2 [2009] EWCA Crim 1251, [3].

4.32 It was assumed that the content described feedback on the quality of the drugs purported to have been supplied. Professor Ormerod considered the decision by the Court of Appeal to be incorrect because the Crown did not rely on the content of the text messages for the truth of whether the quality was good or bad, or the nature of what had been supplied. The issue was whether the appellant had supplied a controlled drug, not the quality of the drugs supplied, which was irrelevant.[1] This is undoubtedly correct, and the argument illustrates the absurdity of the largely arbitrary line between hearsay and non-hearsay statements. To conclude that anything inferred from a

statement is not hearsay whereas anything directly stated is to establish a distinction that *dances on the head of a pin*. The better approach is to treat all types of assertions – express or intended *and* implied or unintended – as *prima facie* hearsay and leave their admission to the judge on the basis of an analysis of a list of balancing criteria.

1 See *R v MK* [2007] EWCA Crim 3150 where a conversation over a telephone by covert recording equipment was not considered to be hearsay, and it was therefore admissible without having to comply with the statutory provisions relating to hearsay.

4.33 This point is illustrated in the next case. In *R v Twist*,[1] the issue was the admissibility of text messages sent over mobile telephones. Whether the text messages were admissible depends on the 'matter stated', which will usually be a fact, but may also be an opinion in accordance with s 115(2) of the Criminal Justice Act 2003. In determining the general approach to take whether the hearsay rules apply in this way, Hughes LJ set out the following approach:

> i) identify what relevant fact (matter) [the statement] is sought to prove;[2]
>
> ii) ask whether there is a statement of **that matter** in the communication. If no, then no question of hearsay arises (whatever other matters may be contained in the communication);
>
> iii) If yes, ask whether it was one of the purposes (not necessarily the only or dominant purpose) of the maker of the communication that the recipient, or any other person, should believe **that matter** or act upon it as true? If yes, it is hearsay. If no, it is not.'[3] (emphasis in the original)

1 [2011] EWCA Crim 1143, [2011] Crim LR 793 (note); note the criticism of Hirst, 'Hearsay, confessions and mobile telephones' 491–3.
2 Hughes LJ indicated at [11] that it must be a relevant matter.
3 [2011] EWCA Crim 1143, [17].

4.34 Hughes LJ went on, at [18], to indicate that the '… answers to these questions will be case-sensitive. The same communication may sometimes be hearsay and sometimes not, depending on the matter for which it is relied upon and the fact which it is sought to prove.'[1] While correct, this line of argument emphasizes the largely arbitrary nature of the distinction. A text message commenting on the quality of drugs bought will not be a hearsay statement and can be adduced in support of a contention that the recipient actually sold drugs. However, a statement to the effect, 'thanks for selling me those drugs' will be inadmissible hearsay. And an argument might be made that what was sought to be established was the state of mind of the maker of the message, not whether drugs were actually sold by the recipient of the message. And that may be admissible, depending on the issue to be proved.

1 Note the criticism by Hirst, 'Hearsay, Confessions and Mobile Telephones' 491–2.

4.35 Evidence of the actions of others recorded in digital form is certainly hearsay. In the Australian case of *Hansen Beverage Company v Bickfords (Australia) Pty Ltd*,[1] working television sets in homes were monitored by a meter system that recorded that a person was physically located in the home when he registered his presence by pressing a button when a television was on. This was for the purposes of establishing the size of the audience that might be watching a particular programme. That the evidence was produced on a print-out and was automatically recorded by software was not at issue. Middleton J, it is suggested correctly, identified the evidence as

hearsay because it was a representation of fact that a certain number of people clicked on the buttons. The judge commented:

> Undoubtedly, Hansen seeks to prove the estimated audience sizes for a particular program derived by statistical methods from the data, but such data is not automatically recorded by the meters without the human intervention of deliberately pressing the button to show a person or persons are in the room where the television is on. When the people are in the room they intend to, and do, make the representation to assert the existence of this fact, the existence of which needs to be proved to form the basis of the statistical analysis. It seems to me that the necessary reliance by Hansen on the data derived from the sample homes must involve the representation … by a person that the person was in the room on the relevant occasion, namely when the television is operating.[2]

1 [2008] FCA 406.
2 [2008] FCA 406, [125].

Business and other documents

4.36 In this regard, it is useful to review the cases that considered s 24 of the Criminal Justice Act 1988. This provision, which provided for the admission in criminal proceedings of business and other documents, was the predecessor provision to s 117 of the Criminal Justice Act 2003. In *Brown v Secretary of State for Social Security*,[1] the Secretary of State adduced evidence of statements from computer records by way of two witnesses where the identity of the persons who supplied the information could not be established. It was submitted on behalf of the appellant that the two statements were inadmissible because they did not comply with the terms of s 24. Section 24 was written to enable business documents to be admissible without the need to call the maker where the documents formed part of records which the maker could not be expected to know anything about in detail, and which were created in the course of trade or business. The members of the Divisions Court, Balcombe LJ and Collins J, agreed that the statements were not admissible under s 24(4) of the Criminal Justice Act 1988, 'as there was no evidence that it was impossible that the makers of the statements would have no recollection of the matters referred to in their statements'.[2] In comparison, the members of the Court of Appeal (Criminal Division) in the case of *R v Derodra*[3] rightly, it is suggested, admitted the contents of a police 'CRIS' report, which was a computerized record of incidents of crime under s 24. In this instance, the person who reported the crime to the police – the lodger of the appellant – could not be found to give evidence of his complaint. It was the statement of the lodger that was to be relied upon testimonially, not that of the police officer who made the relevant entry.[4]

1 [1995] COD 260 (DC).
2 [1995] COD 260 (DC), 262.
3 [2000] 1 Cr App R 41 (CA), [1999] Crim LR 978 (note).
4 For a commentary and references to relevant article, see '*R v Derodra*' [1999] Crim LR 978 (note).

4.37 In *Vehicle and Operator Services Agency v George Jenkins Transport Limited*,[1] the prosecution had to prove that certain commercial drivers had failed to properly record their journeys with the tachographs in their vehicles, and had worked beyond the number of hours that were permitted without the prescribed rest periods or breaks. To discharge this burden, the prosecution sought to put in evidence a number of drivers' time sheets pursuant to s 24. On a preliminary point, the trial judge ruled

them inadmissible and dismissed all charges against the defendants. The prosecutor appealed, and the appeal raised a number of issues regarding the interpretation of these provisions. First, the provisions in s 24, described by Mackay J as 'criteria or gateway' provisions,[2] must be satisfied before the second issue is addressed, that is whether the documents in question can be admitted in evidence. Mackay J quoted[3] from the judgment of Roch LJ in *R v Foxley (Gordon)*:[4]

> Section 24 deals with the statements in a document and makes such statements admissible of any fact of which direct oral evidence would be admissible if two conditions are satisfied. The wording of condition (ii) demonstrates that Parliament anticipated that courts would draw inferences as to the personal knowledge of the person supplying the information of the matters dealt with. The purpose of section 24 is to enable the document to speak for itself; the safeguard being the two conditions and the other statutory provisions applicable, for example in the case of a statement made for the purpose of a criminal investigation, one of the requirements of section 23(2) or the requirements of section 23(3) have to be fulfilled.[5]

1 [2003] EWHC 2879 (Admin).
2 [2003] EWHC 2879 (Admin), [10].
3 [2003] EWHC 2879 (Admin), [24].
4 [1995] 2 Cr App Rep 523 (CA), [1995] Crim LR 636 (note).
5 [1995] 2 Cr App Rep 523 (CA), 536.

4.38 In this instance, Mackay J and Kennedy LJ agreed that the documents satisfied the criteria provisions, and were admissible and self-proving in evidence.[1] Kennedy LJ also noted the criticisms that Professor Smith made of the decision in *R v Foxley (Gordon)*,[2] although it was observed that a further analysis of another case[3] by Professor Smith was capable, if it was adjusted slightly, of applying to the case in hand.[4]

1 [2003] EWHC 2879 (Admin), [30], [34].
2 [1995] 2 Cr App Rep 523 (CA), J C Smith [1995] Crim LR 636 (note).
3 *R v Ilyas and Knight* [1996] Crim LR 810, 811–12.
4 [2003] EWHC 2879 (Admin), [34].

4.39 Section 24 is succeeded by s 117 of the Criminal Justice Act 2003. Section 117(1) to (5) read:

> Business and other documents
>
> (1) In criminal proceedings a statement contained in a document is admissible as evidence of any matter stated if—
>
> (a) oral evidence given in the proceedings would be admissible as evidence of that matter,
>
> (b) the requirements of subsection (2) are satisfied, and
>
> (c) the requirements of subsection (5) are satisfied, in a case where subsection (4) requires them to be.
>
> (2) The requirements of this subsection are satisfied if—
>
> (a) the document or the part containing the statement was created or received by a person in the course of a trade, business, profession or other occupation, or as the holder of a paid or unpaid office,
>
> (b) the person who supplied the information contained in the statement (the relevant person) had or may reasonably be supposed to have had personal knowledge of the matters dealt with, and

> (c) each person (if any) through whom the information was supplied from the relevant person to the person mentioned in paragraph (a) received the information in the course of a trade, business, profession or other occupation, or as the holder of a paid or unpaid office.
>
> (3) The persons mentioned in paragraphs (a) and (b) of subsection (2) may be the same person.
>
> (4) The additional requirements of subsection (5) must be satisfied if the statement—
>
> (a) was prepared for the purposes of pending or contemplated criminal proceedings, or for a criminal investigation, but
>
> (b) was not obtained pursuant to a request under section 7 of the Crime (International Co-operation) Act 2003 (c. 32) or an order under paragraph 6 of Schedule 13 to the Criminal Justice Act 1988 (c. 33) (which relate to overseas evidence) .
>
> (5) The requirements of this subsection are satisfied if—
>
> (a) any of the five conditions mentioned in section 116(2) is satisfied (absence of relevant person etc), or
>
> (b) the relevant person cannot reasonably be expected to have any recollection of the matters dealt with in the statement (having regard to the length of time since he supplied the information and all other circumstances).

4.40 The provisions of s 117, dealing with the business document exception, are very wide and permit the admission into evidence of multiple hearsay,[1] although the various foundational conditions set out in s 117 must be satisfied. In *R v Humphris*[2] the Crown sought to adduce evidence of the appellant's previous convictions under s 117. For that purpose, they relied on a statement of officer Grimes, who retrieved relevant records from Essex Police computer facility, the contents of which were in turn derived from staff of the Essex Police Force, who acted under a duty to record information and who either had or may reasonably be supposed to have had personal knowledge of the matters dealt with in the records. These records were attached to officer Grimes' statement. Section 117 provides certain conditions that must be fulfilled before evidence can be admitted. The defence accepted that the provisions of s 117(2)(a) were complied with, but argued that for each record of the appellant's previous conviction, s 117(2)(b) required the statement to have been obtained from each complainant as the relevant person, rather than the police officer who actually recorded the information. Although Lord Woolfe upheld the conviction of the appellant, he agreed and held that the necessary foundations for the admissibility of the evidence were not properly laid.

1 A point made by Professor Tapper, when he indicated that some electronic information will be collated from other statements, thus constituting multiple hearsay: Colin Tapper, 'Electronic evidence and the Criminal Justice Act 2003' (2004) 10 CTLR 161; an example would be proving the links of the continuity of evidence between the withdrawal of cash from an ATM to demonstrating the entering of the transaction in the customer's account.

2 [2005] EWCA Crim 2030; for a similar point, also see *Maher v DPP* [2006] EWHC 1271 (Admin).

4.41 Where a document is put in under the provisions of s 114 and s 117, care must be taken over any content that is hearsay.[1] In addition, the trial judge must ensure that the members of the jury understand the purpose of admitting the document. In *R v Horncastle*,[2] there was an email statement made by an ISP which identified the

appellant and his address as the possible holder of an email account suspected to have been used to send abusive images of children. The ISP acknowledged that this information could have been supplied by the email account holder impersonating the appellant. The prosecution adduced this email to show the address of the place (the appellant's home) where the police raid took place, but not to prove the fact that the account was that of the appellant or used by the appellant. (In fact, no evidence of abusive images of children was found on the appellant's computer, although there was evidence that the appellant's lodger had used the email account.) No directions were given by the trial judge as to the limited purpose for which the ISP's email was adduced. On appeal, Thomas LJ held that the judge's failure to explain the use was a material misdirection, as the jury could have used the ISP's email to link the appellant to the email account. The appellant's appeal was allowed and his conviction was set aside.

1 Where a print-out from the Police National Computer was correctly admitted into evidence, all of the conditions under s 117 having been met, see *R (on the application of Wellington) v DPP* [2007] EWHC 1061 (Admin).

2 [2009] EWCA Crim 964; note also *DPP v Leigh* [2010] EWHC 345 (Admin), where the prosecution did not rely on a record for the purpose of establishing the veracity of any of the matters recorded.

Judicial discretion to exclude

4.42 A trial judge also has the ability to refuse to admit a statement in accordance with s 126(1)(b) where 'the court is satisfied that the case for excluding the statement, taking account of the danger that to admit it would result in undue waste of time, substantially outweighs the case for admitting it, taking account of the value of the evidence'. A similar provision exists as a component of s 8 of New Zealand's Evidence Act 2006. Section 8 of the New Zealand Evidence Act provides:

> In any proceeding, the Judge must exclude evidence if its probative value is outweighed by the risk that the evidence will—
>
> (a) have an unfairly prejudicial effect on the proceeding; or
>
> (b) needlessly prolong the proceeding.
>
> (2) In determining whether the probative value of evidence is outweighed by the risk that the evidence will have an unfairly prejudicial effect on a criminal proceeding, the Judge must take into account the right of the defendant to offer an effective defence.

Concluding observations

4.43 Almost everybody now uses digital data, whether their interaction is by way of the ether – a terminal linked by software to a server located in an unknown location – or from a physical device. Software code has become part of the everyday fabric of the majority of people. This means we are all, wittingly or unwittingly, assessing digital evidence every day: from whether to trust that incoming email from an unknown source, to dealing with the veracity of content from networking sites.

4.44 The digital world is now awash with evidence: direct statements over the Internet; communications between telephones and other devices; messages made by

a known author, anonymously or by somebody that cannot be traced. Every day we are dealing with the multiplicity of direct and indirect assertions (whether factually accurate or not), in the form of statements by one person or relayed, correctly or incorrectly, by others, and the interplay between them and the reality of the physical world. For the first time, we are all assessing evidence every day.

4.45 It cannot be beyond the ability of lawyers to distinguish the various components of language and communications during a trial to test the evidence effectively without complex rules on hearsay.

5

Software code as the witness

Stephen Mason

5.1 The aim of this chapter is to illustrate how software code can affect the examination and introduction of electronic evidence in legal proceedings. The topic is considered in the context of software code as the 'witness'. It is important to understand how software can affect an assessment of the truth in any given set of facts. Failure to appreciate this can lead to unfairness in legal proceedings and incorrect decisions.

5.2 A digital computer is like a mechanical device, where switches replace gears, and the switches are miniaturised. However, it is impossible to build a mechanical device that reflects the functionality of a modern digital computer, because such a device would require both a machine built on a colossal scale and the use of materials beyond the strengths or machine tolerances of what is possible to mechanically manufacture. To complete the picture, physical digital devices, as indicated in chapter 1, cannot work without the software written by programmers and the input by users.

5.3 It follows that electronic evidence could be treated as a joint statement that is:

> (i) partly made by the person inputting data (such as typing an email or word document, inserting a PIN, filling in forms over the internet – in essence anything a person does when interacting with a devices), and
>
> (ii) partly made by the hundreds of programmers who are responsible for writing the software that produces the data.

5.4 For this reason, there is an argument, as proposed by Steven W. Teppler,[1] that all forms of evidence in digital form remain hearsay, because software code conveys information.[2] Teppler gives the example of United States Patent Office Number 5,619,571, which includes some uncompiled source code that contains the following lines of code in the application:

```
ptrFIXUP   fixupBase  =  NULL;  //  Base pointer for fixups

ptrFIXUP   fixupMap  =  NULL;  //  pointer used to 'walk off of base'

FIXUP      IVFixup;  //  ISII  Verification fixup

memset(&IVFixup,0,sizeof(FIXUP));

// Allocate a buffer to build the IFD  (If this fails, we are F'd)[3]
```

1 Stephen W Teppler, 'Digital data as hearsay' (2009) 6 *Digital Evidence and Electronic Signature Law Review* 7.

2 Stephen W Teppler, 'Testable reliability: a modernized approach to ESI admissibility' (2014) 12 *Ave Maria Law Review* 213, 255.

3 U.S. Patent No. 5,619,571 (issued Apr. 8, 1997), 17–18, lines 10–14.

5.5 What this comment indicates is an acknowledgment of the possibility of a weakness in the software code that has been written, not that the software code is

Stephen Mason, 'Software code as the witness', in Stephen Mason and Daniel Seng (eds.), *Electronic Evidence* (4th edn, University of London 2017) 88–100.

or will be at fault. In this regard, it is useful to understand more fully the nature of source code. For instance, Svein Willassen explains the complex nature of software as follows:[1]

> Software is written as source code. The source code is written by the programmer, by entering instructions in an editor. The sequence of instructions defines the function of the program, such as taking input from the user, performing calculations, showing output on the screen and so on. This source code is then usually compiled into an executable program (an executable file causes a computer to perform tasks in accordance with the instructions), which is distributed to the users of the program. The source code cannot be derived completely from the executable program.

1 Svein Yngvar Willassen, 'Line based hash analysis of source code infringement' (2009) 6 Digital Evidence and Electronic Signature Law Review 210.

5.6 In *Computer Edge Pty Limited v Apple Computer Inc*, Gibbs CJ offered the following explanation of the various parts of a computer program:

> A computer program is a set of instructions designed to cause a computer to perform a particular function or to produce a particular result. A program is usually developed in a number of stages. First, the sequence of operations which the computer will be required to perform is commonly written out in ordinary language, with the help, if necessary, of mathematical formulae and of a flow chart and diagram representing the procedure. In the present case if any writing in ordinary language (other than the comments and labels mentioned below) was produced in the production of Applesoft and Autostart, no question now arises concerning it. Next there is prepared what is called a source program. The instructions are now expressed in a computer language—either in a source code (which is not far removed from ordinary language, and is hence called a high level language) or in an assembly code (a low level language, which is further removed from ordinary language than a source code), or successively in both. Sometimes the expression 'source code' seems to be used to include both high level and low level language. In the present case, the source programs were written in an assembly code, comprising four elements, *viz.*:
>
> > (a) labels identifying particular parts of the program;
> >
> > (b) mnemonics each consisting of three letters of the alphabet and corresponding to a particular operation expressed in 6502 Assembly Code (the code used);
> >
> > (c) mnemonics identifying the register in the microprocessor and/or the number of instructions in the program to which the operation referred to in (b) related; and
> >
> > (d) comments intended to explain the function of the particular part of the program for the benefit of a human reader of the program.
>
> The writing has been destroyed, although it is possible to reconstruct the mnemonics, but not the labels and comments, which were comprised in it.
>
> The source code or assembly code cannot be used directly in the computer, and must be converted into an object code, which is 'machine readable', *i.e.* which can be directly used in the computer. The conversion is effected by a computer, itself properly programmed. The program in object code, the object program, in the first instance consists of a sequence of electrical impulses which are often first stored on a magnetic disk or tape, and which may be stored permanently in a ROM ('read only memory'), a silicon chip which contains thousands of connected

> electrical circuits. The object code is embodied in the ROM in such a way that when the ROM is installed in the computer and electrical power is applied, there is generated the sequence of electrical impulses which cause the computer to take the action which the program is designed to achieve. The pattern of the circuits in the ROM may possibly be discerned with the aid of an electron microscope but it cannot be seen by the naked eye. Obviously, the electrical impulses themselves cannot be perceived. However the sequence of electrical impulses may be described either in binary notation (using the symbols 0 and 1) or in hexadecimal notation (using the numbers 0-9 and the letters A-F), and it is possible to display the description on the visual display unit of the computer, and to print it out on paper. And, as has been said, it is also possible to reconstruct the mnemonics in the source code. It will have been seen from this account that a program exists successively in source code and in object code, but the object code need not be written out in binary or hexadecimal notation in the process of producing and storing the program.[1]

1 [1986] FSR 537, 541–2.

5.7 The term 'source code' is also the subject of a commentary by Jacob J in the case of *Ibcos Computers Ltd v Barclays Mercantile Highland Finance Ltd*:

> The program the human writes is called the 'source code'. After it is written it is processed by a program called a compiler into binary code. That is what the computer uses. All the words and algebraic symbols become binary numbers. Now when a human writes he often needs to make notes to remind himself of what he has done and to indicate where the important bits are. This is true of life generally and for programmers. So it is possible to insert messages in a source code. A reader who has access to it can then understand, or understand more readily, what it going on. Such notes, which form no part of the program so far as the computer is concerned, are called 'comments'. They are a kind of side-note for humans. In the DIBOL and DBL programs with which I am concerned, a line or part of a line of program which is preceded by a semi-colon is taken by the complier as a comment. That line is not translated by the compiler into machine code. The program would work without the comment. It follows that although computers are unforgiving as to spelling in their programs, they do not care about misspelt comments in the source code. If a line of operational code (a 'command line') is modified by putting a semi-colon in front of it, it ceases to be operational. The computer treats the code as a mere comment. Computer programmers sometimes do this with a line which pre-exists when they no longer want that line, but are not sure they may not need it in the future. Or, if the programmer thinks he may want to add a feature to his program in the future he may put in a comment allowing for this. He is unlikely in the latter instance to put in detailed code only to comment it out. A general note will do.
>
> Source code, being what humans can understand, is very important to anyone who wants to copy a program with modifications, for instance to upgrade it. It is the source code which shows the human how it all works, and he or she will also get the benefit of all the comments laid down by the original programmer. Software houses not surprisingly normally keep their source code to themselves and confidential.[1]

1 [1994] FSR 275, 286.

5.8 There is a distinction between the code written by programmers that provides instructions to the computer, and the comments made by the programmer writing the code. If the software code is inaccurate, or if an instruction written by a programmer

acts on information or a further instruction that is incorrect, then the code will probably fail to instruct the computer in the way the programmer intended. However, comments by a programmer that do not form part of the instructions cannot necessarily be considered to be part of the code.

The classification of digital data

5.9 The starting point to this analysis is an attempt at classifying software code as digital data. To this end, Professor Ormerod, the commentator in a report on the case of *R v Skinner*,[1] suggested there were three questions to consider for every type of digital data:

> (i) Who or what made the representation.
>
> (ii) Whether the representation was hearsay or not.
>
> (iii) Whether the evidence is authentic.[2]

1 [2005] EWCA Crim 1439.

2 David C Ormerod, 'Evidence: information copied from one website to another' [2006] Crim LR 56.

5.10 In *Elf Caledonia Ltd v London Bridge Engineering Ltd*, Lord Caplan noted the following:

> The defenders suggested that there are three categories of use for computers. They can be used to record data without the need of human intervention. The Spectra-Tek programme was described as being of this type. It was said that what this programme prints out may be regarded as real evidence. However Counsel had to concede that even this type of computer exercise depends on the reliability of the material programme. Unless it is properly programmed it will not store and regurgitate facts accurately. …
>
> Another category of computer use was said to be where data is recorded by the computer and the data is put in manually. Thus Piper would regularly send information to the beach and this would be entered in the computer system. It was accepted that to prove this material would involve some hearsay evidence unless the persons who entered the material in the computer were led as witnesses. However the defenders did not explore just what evidence would be required in the situation under consideration. In general it seems to me that there must be many cases where it would not be practicable to lead the person who generated the data and the person who fed it into the computer so that there must be some practical limits as to what proof can be expected in this kind of computer evidence.
>
> It was submitted that the third type of computer situation is where the computer is used by experts to carry out calculations or simulations. It was claimed that in this kind of situation the general rules relating to expert evidence should be applied. Certainly in this kind of situation one can get a distorted result if one factor is in-putted wrongly. The kind of computer models used by experts of course generally requires more than normal discrimination and judgment in the selection of in-put material. Thus the expert will have to prove how the input material was arrived at and the justification for selecting what was put in. However I am not sure that the three categories of computer exercise referred to by the defenders' Counsel can be distinguished quite as neatly as he attempts. Even in a simple office system distorted results will arise if the proper material is not fed into the computer. Thus it was argued that the first requirement in

> considering computer evidence given by an expert is to consider the input. That may be so but it cannot be exclusive to expert computer evidence. Of course it was said that the best evidence of in-put and out-put material is in the print-outs of such material.[1]

1 [1997] ScotCS 1, 898–900, sub nom *Elf Enterprise Caledonia Ltd v London Bridge Engineering Limited* [1997] ScotCS 1, 2.

5.11 Based on this categorization, Professor Ormerod noted that some types of computer-generated representations do not infringe the hearsay rule.[1] If a computer carries out the instructions of the program that has been written by humans to create such data, it may be right to suggest that such data are probably accurate without the need to test whether they are correct. But if the time as noted by a clock on a camera linked to an ATM is to be offered into evidence to link the accused to the murder of the person whose card was used in the ATM, then the time as data will have to be adduced as to its truth, as in the case of *Liser v Smith*,[2] and there will be a need to validate the clock, and verify the time and date set by a human being.[3]

1 Although he accepts that s 129 of the Criminal Justice Act 2003 may need to be considered. For a commentary on s 129, see John R Spencer, *Hearsay Evidence in Criminal Proceedings* (Hart 2008) ch 3.
2 254 F.Supp.2d 89 (D.D.C. 2003).
3 As noted by Colin Tapper: 'Reform on the law of evidence in relation to the output from computers' (1995) 3 Intl J L & Info Tech 79, 85 fn 44.

5.12 To the same end, Professor Smith distinguished between the types of representations that the code in a device can make,[1] and argued that where the computer is instructed to perform certain functions, many of which are performed in a mechanical way (such as the addition of the time and date on an email), in such circumstances the computer is producing real evidence, not hearsay. In illustrating the point he was making, Professor Smith gave a number of examples where evidence is not hearsay.[2] One example was that of Six's thermometer (commonly known as a maximum minimum thermometer), which he referred to as an instrument and not a machine. This is correct. The thermometer provides three readings: the current temperature, and the highest and the lowest temperatures reached since it was last reset. A human being can give evidence of his observation of the precise location of the mercury against the scale at a given time and date. The witness might be challenged as to the truthfulness of his recollection without calling into question the accuracy of the instrument. Such evidence will not be hearsay. Alternatively, the precision of the scale on the thermometer might be open to scrutiny, in which case it will be necessary to have the instrument tested by an appropriately qualified expert.[3]

1 J C Smith, 'The admissibility of statements by computer' [1981] Crim LR 387.
2 Smith, 'The admissibility of statements by computer' 387, 390.
3 This was also discussed by Penelope A Pengilley, 'Machine information: is it hearsay?' (1982) 13 Melbourne University Law Review 617, 625.

5.13 Further examples considered by Professor Smith included a camera that records an image, a tape recorder that records sound, and a radar speedmeter that records the speed of a vehicle. In 1981, each of these machines was mechanical in construction, with the exception of the radar speedmeter, which also incorporated components that were instruments. None of the examples involved devices controlled by software written by human beings. Although it is possible to alter the image from a camera or

the sound from a tape recording, or for a human being to lie about the reading from a radar speedmeter, nevertheless the evidence from such devices would not be hearsay.

5.14 In respect of software, Professor Smith indicated that a programmer may make mistakes (errors are common, for which see the chapter on 'reliability'), but mistakes can also be made when deciding the scale on a thermometer. He went on to suggest that '[t]his consideration goes to weight rather than admissibility. In any event it certainly has nothing to do with the hearsay rule'.[1]

1 Smith, 'The admissibility of statements by computer' 387, 390. One answer to this issue has been proposed by Professor Pattenden – that s 129(1) of the Criminal Justice Act 2003 be replaced 'with a single test of admissibility for all factual representations that are not in substance the statement of a person but "machinespeak", that is, those whose content is the outcome of creating machine-processing', for which see Rosemary Pattenden, 'Machinespeak: section 129 of the Criminal Justice Act 2003' [2010] Crim LR 623, 636–7; Professor Pattenden discusses the conflicting opinions relating to s 129(1) in detail.

5.15 Professor Seng proposed an analysis in 1997:

> Computers which are used as data processing devices can be classified into the following categories: devices which accept human-supplied input and produce output, self-contained data processing devices which obtain input or take recordings from the environment without human intervention, and a hybrid of the two.[1]

1 Daniel Seng, 'Computer output as evidence' (1997) 130 Sing JLS 173.

5.16 Steven Teppler also accepted that it is possible to categorize data into three, treating digital data as hearsay:

> (i) The memorandum 'created' by a human.
>
> (ii) Digital data generated in part with human assistance.
>
> (iii) Digital data generated without a human being.[1]

1 Teppler, 'Testable reliability', 235–40.

5.17 Teppler has also suggested that a 'fourth potential category, for which there has been no judicial analysis, has recently emerged as a consequence of computer programs that "listen and respond" to questions in natural language and with a "voice" that closely mimics a "real" human'.[1]

1 Teppler, 'Testable reliability', 235.

5.18 The authors of *Archbold* have also divided digital data into three categories:

> (i) Where the device is used as a processor of data.
>
> (ii) Where the software records data where there is no human input involved.
>
> (ii) Where there is data recorded and processed by software that has been entered by a person, directly or indirectly.[1]

1 James Richardson (ed), *Archbold: Criminal Pleading, Evidence and Practice 2016* (64th rev edn, Sweet & Maxwell 2016) paras 9-11–9-14.

5.19 It is proposed that the three categories outlined by Professor Seng, Steven Teppler and the authors of *Archbold* be slightly amended to read as follows:

(i) Content written by one or more people (that is, where the device is used as a processor of data).

(ii) Records generated by the software that have not had any input from a human.

(iii) Records comprising a mix of human input and calculations generated by software.

Each of these categories is discussed below.

Content written by one or more people

5.20 Records of electronic content that are written by one or more people include email messages, word processing files and instant messages. Unless the author of the software has included instructions to alter the content of the text that has been typed in by a human, the only function of the device is to store the information that has been input by the human being. However, Teppler suggests that all computer-generated information is hearsay of some sort, and that the data generated by an email program, for instance, remains hearsay because

> the receiving computer is carrying out the stated intent or declaration of some person who instructed the computer to make the assertion on his or her behalf (e.g., a programmer) to carry out some request (and provided that certain conditions are met) that the receiving computer was told by the sending computer as agent for that person, which in turn was requested by a statement or declaration of the person or sender.[1]

1 Teppler, 'Testable reliability' 240.

5.21 Conceptually this must be right, but the status of the instructions issued by the software code at the material time is rarely relevant. This category, artificial as it might appear to be, enables the content that was input by the maker of the statement to be separated from the content made by the author of the software program – in the same way that the printed notepaper with the name of the person or organization, together with other information such as address and telephone number is created by the printer, but is distinct from the content of the letter.

5.22 The content of the software program will not be relevant unless there is a dispute as to what data was entered, when and where it was entered, and by whom. In such circumstances, the relevant witnesses can be called to give oral evidence to determine the truth, failing which a suitably qualified digital evidence practitioner might be called to give evidence about the metadata associated with the document to help ascertain answers to these technical questions.

5.23 By way of example, consider whether a letter typed into a computer is a document produced by a computer. Professor Smith took the view that if the human author printed the document and then read the contents to verify the text, the author authenticates the text. Given this set of facts, the computer is a mere tool. Where the author does not read the print-out, the document remains computer output.[1] Professor Seng suggests that 'it is difficult to see how reading what is clearly a computer-produced document converts it into one not produced by a computer. The print-out remains clearly a document produced by a computer operated as a data storage device'.[2] Professor Smith indicates that the person can authenticate the text after it has

been printed. This does not mean that the act of authentication takes away the fact that the document was created on and remains stored on the device. This distinction can be important, as in the case of electronic wills. The court must establish whether, in the absence of the testator authenticating the will, the testator actually wrote the will and intended it to be their last will and testament. In such cases, it might also be necessary to give consideration to both the content written by the human and the software code that makes up the metadata.[3]

1 *R v Shephard* [1993] Crim LR 295 (note), 297–8.
2 Seng, 'Computer output as evidence' 130, 178 – Professor Seng begins his discussion by asking whether word-processed documents are computer output or recorded computer output: 177.
For cases involving wills in electronic form from Australia, Canada, South Africa and the United States of America, see Stephen Mason, *Electronic Signatures in Law* (4th edn, University of London 2016) paras 10.48–10.66.

5.24 Professor Tapper pointed out that computers include such facilities as spell checkers, calculators and automatic paragraph numbering, amongst other tools. This suggests that a word file (such as a letter) is processed computer output.[1] In his discussion, Professor Smith also discussed the same document being produced by a human typing on a typewriter. If the text – for the sake of illustration, a letter – is written by hand, or typed on a typewriter, or typed into a computer, the resultant content will be the same, other than the type of print, typeface and such like, although the author might cause the data to remain stored on the device if it was a computer.[2] The person writing the letter by hand or on a typewriter might use a dictionary to check their spelling in the same way that spelling can be checked on a computer using the spell checker. Whether the letter is written by hand, typed on a typewriter or on a computer, the letter will then be complete when printed (in the case of the computer) on paper. The method used to record words on paper must be irrelevant, providing that the only evidence to be relied upon is the text that is recorded on the paper. If other factors are in issue, such as the purported author of the document, then clearly an examination of the digital data might be instructive. Professor Seng takes issue with Professor Smith's characterization that the evidential quality of a letter changes immediately when a recipient reads it, without taking into account any characterization of its source. In such a case, where the computer is behaving as a storage device, the rebuttable presumption is that the code operating to make it behave as such is reliable, and issues as to authentication of this code do not enter the evidential analysis, generally speaking. But there can be other software errors, for which see the chapter on the 'reliability' of computers.

1 Colin Tapper, 'Reform of the law of evidence in relation to the output from computers' (1995) 3 International Journal of Law & Information Technology 79, 86–88.
2 A point made by Professor Seng, 'Computer output as evidence' 178.

5.25 The Law Commission in their report[1] noted that the '… present law draws a distinction according to whether the statement consists of, or is based upon, only what the machine itself has observed; or whether it incorporates, or is based upon, information supplied by a human being'.[2] It was further noted that the hearsay rule did not apply to tapes, films or photographs, or to documents produced by machines that automatically record an event or circumstance.[3] This was because the court is not being asked to accept the truth of an assertion made by any person, and the evidence is real evidence, not hearsay.

1 Law Commission, *Evidence in Criminal Proceedings: Hearsay and Related Topics* (Law Com No 245, 1997).
2 Law Commission, *Evidence in Criminal Proceedings: Hearsay and Related Topics*, para 7.43.
3 Law Commission, *Evidence in Criminal Proceedings: Hearsay and Related Topics*, para 7.44.

5.26 That humans generally have control over a computer system is demonstrated in the case of *Ferguson v British Gas Trading Limited*,[1] in which the members of the Court of Appeal rejected arguments submitted that letters sent out automatically by a computer were not the fault of British Gas. Computers only work on instructions given to them, and it followed that a person in British Gas, or authorized by British Gas, must have instructed the computer to initiate the letters in question. In this case, British Gas sent letters to the claimant that the court held were capable of amounting to unlawful harassment contrary to the Protection from Harassment Act 1997. In the words of Jacob LJ: 'British Gas says it has done nothing wrong; that it is perfectly all right for it to treat consumers in this way, at least if it is all just done by computer'.[2] Jacob LJ went on to indicate that he did not follow the reasoning of Martin Porter QC, counsel for British Gas, that '[as] the correspondence was computer generated … [the harassed victim] should not have taken it as seriously as if it had come from an individual'.[3] Jacob LJ noted that computers operate on instructions given to them: '… real people are responsible for programming and entering material into the computer. It is British Gas's system which, at the very least, allowed the impugned conduct to happen'.[4] Likewise, Sedley LJ roundly rejected the pathetic excuse offered by British Gas:

> One excuse which has formed part of British Gas's legal argument for striking out the claim, and which has been advanced as incontestable and decisive, is that a large corporation such as British Gas cannot be legally responsible for mistakes made either by its computerised debt recovery system or by the personnel responsible for programming and operating it. The short answer is that it can be, for reasons explained by Lord Justice Jacob. It would be remarkable if it could not: it would mean that the privilege of incorporation not only shielded its shareholders and directors from personal liability for its debts but protected the company itself from legal liabilities which a natural person cannot evade. That is not what legal personality means.[5]

1 [2009] EWCA Civ 46.
2 [2009] EWCA Civ 46, [5].
3 [2009] EWCA Civ 46, [21].
4 [2009] EWCA Civ 46, [21].
5 [2009] EWCA Civ 46, [51].

Records generated by the software that have not had any input from a human

5.27 Examples of records generated by software controlling a computer without any input from a human include computer data logs for the purposes of tracking activity and diagnostics, number plate recognition,[1] automatic connections made by telephone switches and the records of such calls made for billing purposes,[2] and records of ATM transactions. In one case involving one Antonio Boparan Singh, Singh was convicted of dangerous driving. Part of the evidence adduced by the prosecution included evidence from the event data recorder (EDR) – a device fitted to the airbag system of his vehicle. The EDR established that a force equivalent to 42 mph was lost in one-fifth of a second in the crash. This information helped the police to put Singh's speed at around 72 mph.[3]

1 For judicial consideration of automatic number plate recognition, see *Jackson v R.* [2011] EWCA Crim 1870; *Attorney Generals Reference No 114 – 115 of 2009* [2010] EWCA Crim 1459; *A (Death of a Baby), Re* [2011] EWHC 2754 (Fam); *Najib v R.* [2013] EWCA Crim 86; *Khan v R* [2013] EWCA Crim 2230; *Welsh v R* [2014] EWCA Crim 1027.
2 Rosemary Pattenden, 'Authenticating "things" in English law: principles for adducing tangible evidence in common law jury trials', (2008) 12 E & P 273, suggests that 'self-generated output' can be categorized into two sub-divisions: output that contains no input from human thought, and output that is generated that draws directly or indirectly on information fed into the device by a person: p. 297.
3 Mark Cowan, 'Crime files: picking up the pieces on Midland roads', *Birmingham Mail* (Birmingham, 6 October 2010); an insurance company used data recorded from telematics technology installed in a motor vehicle to disprove 31 claims involving seven accidents over five months: O Ralph, 'Black box data expose £500,000 driver fraud', *Financial Times* (London, 11 June 2016) 4.

5.28 It does not follow that the automatic communications that occur between software code are accurate. For instance, the records from a telephone service provider might be admitted to show that calls were made and received,[1] but it does not follow that the same records can be used as a basis for showing that a SIM card used in a mobile telephone, and purportedly its user,[2] were at a particular location or moved from location to location.[3]

1 For an analysis in the context of New Brunswick, Canada, see *Her Majesty the Queen v Dennis James Oland* 2015 NBQB 244 (third ruling); *Her Majesty the Queen v Dennis James Oland* 2015 NBQB 245 (fourth ruling) and the observations by David M Paciocco, 'Proof and progress: coping with the law of evidence in a technological age' (2013) 11 Canadian Journal of Law and Technology 181, which in turn are disputed in Ken Chasse, 'Guilt by mobile phone tracking shouldn't make "evidence to the contrary" impossible', available at <www.slaw.ca/2016/10/04/guilt-by-mobile-phone-tracking-shouldnt-make-evidence-to-the-contrary-impossible/>.
2 Cell site analysis was the subject of discussion in *Jackson v R.* [2011] EWCA Crim 1870; Reg Coutts and Hugh Selby in their paper 'Safe and unsafe use of mobile phone evidence' (Public Defenders Criminal Law Conference, Sydney, March 2009), <www.publicdefenders.nsw.gov.au/Documents/safeunsafemobilephones.pdf> recommend that defence lawyers pay particular attention to the explanation of cell site analysis set out by Blaxell J in *The State of Western Australia v Coates* [2007] WASC 307, [211]–[220]; Reg Coutts and Hugh Selby, 'Problems with cell phone evidence tendered to 'prove' the location of a person at a point in time' (2016) 13 Digital Evidence and Electronic Signature Law Review 76.
3 Michael Cherry, Edward J Imwinkelried, Manfred Schenk, Aaron Romano, Naomi Fetterman, Nicole Hardin and Arnie Beckman, 'Cell tower junk science' (2012) 95 Judicature 151, 151–52; Aaron Blank, 'The limitations and admissibility of using historical cellular site data to track the location of a cellular phone' (2011) 18 Richmond Journal of Law & Technology 10–12; Herbert B Dixon Jr, 'Scientific fact or junk science? Tracking a cell phone without GPS' (2014) 53 Judges' Journal 37; Graeme Horsman and Lynne R Conniss, 'Investigating evidence of mobile phone usage by drivers in road traffic accidents' (2015) 12 Digital Investigation S30, S37; Alex Biedermann and Joelle Vuille, 'Digital evidence, 'absence' of data and ambiguous patterns of reasoning' (2016) 16 Digital Investigation S86, S94; for the case of Phuong Canh Ngo, see *R v Ngo* [2001] NSWSC 1021 (the sentence); *R v Ngo* [2003] NSWCCA 82 (appeal against conviction); David Patten (Judicial Officer Conducting Inquiry), *Report to the Chief Justice of New South Wales (The Hon J J Spigelman AC) of the Inquiry into the Conviction of Phuong Canh Ngo for the murder of John Newman* (14 April 2009) <www.lawlink.nsw.gov.au/practice_notes/nswsc_pc.nsf/6a64691105a54031ca256880000c25d7/f1ef2541db38ae82ca25759b00052606/$FILE/Report_Phuong_Ngo_140409.pdf>; *Phuong Canh Ngo – Application under Part 7 Crimes (Appeal and Review) Act 2001* [2010] NSWSC 981 (hearing after Report published).

Records comprising a mix of human input and calculations generated by software

5.29 An example of records comprising a mix of human input and calculations generated by software is that of a financial spreadsheet program that contains

human statements (input to the spreadsheet program), and computer processing (mathematical calculations performed by the spreadsheet program). From an evidential point of view, the issue is whether the person or the software created the content of the record, and how much of the content was created by the software and how much by the human. It is possible that the quality of the software acts to undermine the authenticity of the data, which may in turn affect the truth of the statement tendered in evidence. The algorithms in spreadsheet programs are a good example of where the software code affects the truth of the statement. For a more detailed analysis, see the chapter on authentication.

5.30 Professor Pattenden suggests that 'most representations of fact require human intervention at some point',[1] which must be right. The Law Commission report also indicated:

> By contrast, the law does sometimes exclude evidence of a statement generated by a machine, where the statement is based on information fed into the machine by a human being. In such a case, it seems, the statement by the machine is admissible *only* if the facts on which it is based are themselves proved.[2]

1 Pattenden, 'Machinespeak' 623, 633.
2 Law Commission, *Evidence in Criminal Proceedings: Hearsay and Related Topics* para 7.46.

5.31 This comment distinguishes between information fed into a machine (the word 'computer' is not used, but the word 'machine' is presumably meant to include a computer or computer-like device), and the instructions contained in software code written by human beings that are essential for a device to work. Where a person inputs information into a computer, and that information is to be relied upon as to the truth of the statement, then the person should give oral evidence of this action. In contrast, the software code that might be used to transform the raw data into information that can be used is not necessarily relevant, depending on the purpose for which it is adduced in evidence. To this end, the Law Commission[1] compared the cases of *R v Wood (Stanley William)*[2] and *R v Coventry Justices, Ex p Bullard*.[3] In *Wood*, the evidence of the analysis by a computer of tests carried out by chemists was not considered to be hearsay because the chemists gave oral evidence of the results of the tests. The calculations performed by the computer were carried out under the instructions of the person who wrote the software code. The chemists were able to give oral evidence of the results of the tests they performed, but the computer software carried out the actual analysis. The calculations relied upon the software code, which was created by a human being (in this case, a Mr Kellie). The software analysed the data in accordance with the instructions given to it by Mr Kellie. The computer was not capable of analysing the data without the software code. The chemists gave oral evidence of the results of the computer program. This means that the truth of the content of the output of the computer was predicated upon the software code created by Mr Kellie.

1 Law Commission, *Evidence in Criminal Proceedings: Hearsay and Related Topics*, para 7.47.
2 (1983) 76 Cr App R 23 (CA), J C Smith [1982] Crim LR 667 (note).
3 (1992) 95 Cr App R 175 (QB), [1992] RA 79; 'Print-out inadmissible as hearsay' (1993) 57 JCL 232.

5.32 In comparison, the computer print-out in *R v Coventry Justices, Ex p Bullard* included a statement that a person was in arrears with his community charge. This was held to be inadmissible hearsay because the content of the print-out contained information that had been put into the computer by a human, and the print-out had not

been properly proved. The Law Commission, agreeing with the result, would propose a similar analysis as follows:

> An alternative view is that the statement by the machine, properly understood, is conditional on the accuracy of the data on which it is based; and that, if those data are not proved to have been accurate, the statement therefore has no probative value at all. The question of hearsay does not arise, because the statement is simply irrelevant.[1]

1 Law Commission, *Evidence in Criminal Proceedings: Hearsay and Related Topics*, para 7.48.

5.33 In *Mehesz v Redman*,[1] Zelling J concluded that the output of an auto-lab data analyser was hearsay, given that the analysis relied on software where the writer of the software had not been called, and where modifications had been made but the person responsible for the modifications had not been called either. A similar decision was made in *Holt v Auckland City Council*,[2] where evidence of the analysis of the amount of alcohol in a blood sample was excluded by the New Zealand Court of Appeal because the truth of the statement tendered was predicated upon the software code written by a programmer who was not called to give evidence, which meant there was a gap in the continuity of proof. In contrast, in *Wood*, the oral evidence of the results of the tests were read out by the chemists from print-outs of the computer (which was real evidence), and if the results were to be challenged for their accuracy, then the integrity of the software program might need to be tested.

1 (1979) 21 SASR 569.
2 [1980] 2 NZLR 124.

5.34 The instructions written by a human in the form of software code can, depending on the circumstances, be just that: instructions to the machine to perform a particular task. This is illustrated in the case of *Maynard*.[1] An item of software, called a trace, had been written to ascertain whether a particular employee was obtaining access to private information in a computer system, and if so, to record the time and date that the employee viewed the data. The employee was subsequently prosecuted. The magistrate refused to admit the evidence of the print-out of the trace data, partly because he considered the record of the time and date to be hearsay. On appeal, Wright J rejected this analysis. The person that wrote the code gave evidence at trial, both as to the reason for writing the code and as to how it worked. The judge indicated that:

> … it seems to me that once the trace was applied to the respondent's log-on identification, the process then undertaken by the trace was entirely mechanical in that the peregrinations through the database by that computer user was automatically traced through the system and were recorded and stored ready for retrieval in report form as soon as the trace print-out was called for.[2]

1 (1993) 70 A Crim R 133, sub nom *Rook v Maynard* (1993) 126 ALR 150.
2 (1993) 70 A Crim R 133, 141.

5.35 Wright J then went on to illustrate the separate steps:

> Although much more complex in its operation than the following description suggests, the process, stripped to its essentials, involved (a) The implementation of the trace program and its attachment to the respondent's log-on identification. This was a human function proved by direct evidence from Mr Poulter [the person who wrote the code]. (b) Once attached, the trace followed the log-on

> identification number and the user and (c) when the user tapped into or called up a particular file from the database, the trace was able to store details of this event in its memory for subsequent retrieval.

5.36 There was no evidence that suggested that the trace program modified any other programs in the computer, and if there were any such failings, the program designer could have been cross-examined on them. For this reason, the statement was not hearsay.

Challenging the code to challenge the truth of the statement

5.37 One of the most frequently mounted challenges with evidence in digital form is the admissibility of the output from breath-testing devices. Such challenges are attempted across jurisdictions, but the legislation put in place usually provides that where a device is authorized by an appropriate authority, judges do not have the power to require the prosecution to reveal the software code. However, in *State of New Jersey v Chun*, the Supreme Court in New Jersey in the United States ordered the software of a new breath-testing device – the Alcotest 7100 MK111-C – to be reviewed in detail and tested for scientific validity.[1] After extensive testing, the court concluded that the Alcotest, using New Jersey Firmware version 3.11, 'is generally scientifically reliable', but ordered modifications to enable its results to be admitted into legal proceedings.[2] The analysis of the source code indicated that there was a fault when a third breath sample was taken, which could cause the reading to be incorrect, and the court saw fit to order a change in one of the formulae used in the software. This is a significant decision because the court accepted, albeit implicitly, that the software that controlled the device, written by a human, was defective. This in turn meant that had the code not been remedied, the data relied upon for the truth of the statement would be defective and therefore this would affect the accuracy and truthfulness of the evidence.

1 194 N.J. 54, 943 A.2d 114; an application authorising the discovery of source code used in the Intoxilyzer 5000 breath test equipment failed for procedural reasons in *State of Florida v Bjorkland* 924 So.2d 971 (Fla. 2d DCA 2006).

2 943 A.2d 114, 120.

6

The presumption that computers are 'reliable'

Stephen Mason

6.1 This chapter considers the common law presumption in the law of England and Wales that 'In the absence of evidence to the contrary, the courts will presume that mechanical instruments were in order at the material time'. The concept of 'judicial notice' is also considered in this chapter.[1] The Law Commission formulated this presumption in 1997.[2] The reasons given by the Law Commission for the introduction of this presumption make it clear that the words 'mechanical instruments' include computers and computer-like devices – even though computers and computer-like devices are not mechanical instruments. Second, judges have, although not exclusively, used the term 'reliable' in relation to computers. The purpose of this chapter is to consider the introduction of a presumption of 'in order' or 'reliability' or 'working properly' in relation to mechanical instruments generally, and to explain why the term 'reliable' in relation to computers and computer-like devices is not accurate. It must be emphasized that the examples of the failure of computers and similar devices discussed in this chapter are provided to demonstrate the problems that occur, and do not represent the totality of illustrations that could be used, nor the volume of errors that have occurred or will occur in the future. It is suggested that judicial notice be taken of these examples, especially because they contradict the presumption.

1 *Halsbury's Laws* (5th edn, 2015) vol. 12, paras 712–23.

2 Law Commission, *Evidence in Criminal Proceedings: Hearsay and Related Topics* (Law Com No 245, 1997), para 13.13; for the United States of America, see Coleen M Barger, 'Challenging judicial notice of facts on the internet under Federal Rule of Evidence 201' (2013) 48 *University of San Francisco Law Review* 43.

The purpose of a presumption

6.2 The aim of a presumption, which allocates the burden of proof,[1] is to alleviate the need to prove every item of evidence adduced in court, or to reduce the need for evidence in relation to some issues, to save 'the time and expense of proving the obvious'.[2] In an appeal before the Supreme Court of South Australia in the case of *Barker v Fauser*[3] regarding the accuracy of the readings of a weighbridge, Travers J explained the rationale as follows:

> It is rather a matter of the application of the ordinary principles of circumstantial evidence. In my opinion such instruments can merely provide prima-facie evidence in the sense indicated by *May v. O'Sullivan* [(1955) 92 CLR 654]. They do not transfer any onus of proof to one who disputes them, though they may, and often do, create a case to answer. Circumstantial evidence is something which is largely based upon our ordinary experience of life. … It is merely an application of this principle to our ordinary experience in life which tells us of the general probability of the substantial correctness of watches, weighbridges and other such instruments. If they are instruments or machines of a type which we

> know to be in common use our experience tells us that this is suggestive of their substantial correctness. Experience also tells us that they are rarely completely accurate, but usually so substantially accurate that people go on using them, and that subject to a certain amount of allowance for some measure of incorrectness, they act upon them.[4]

1 Colin Tapper, *Cross and Tapper on Evidence* (12th edn, Oxford University Press 2010) 131.
2 *Holt v Auckland City Council* [1980] 2 NZLR 124, per Richardson J at 128.
3 (1962) SASR 176.
4 (1962) SASR 176, 178–179.

6.3 This explanation justifies the rationale for the presumption that mechanical instruments were in order at the material time. However, it appears that the presumption exists on the basis of expediency. In admitting evidence from a mechanical instrument or similar device, judges have not justified the presumption on scientific evidence, but have substituted for it concepts such as 'common use', 'ordinary experience' or 'substantial correctness'.

6.4 Consider the accuracy of a watch – just because a watch has passed tests of accuracy at one moment in time does not preclude its mechanical parts from failing subsequently. In *Barker v Fauser* Travers J put the discussion of the accuracy of mechanical instruments into its overall context as follows:

> My view on the subject of such instruments is that reliance on them is basically an application of circumstantial evidence. The fact that people go on relying upon watches, speedometers, or even hearing aids, seems to be some circumstantial proof that all these things do provide some aid or assistance to those who use them, otherwise they would not go on using them. They are not necessarily accurate, and indeed, probably, most of such instruments on being properly tested would reveal some degree of inaccuracy. But I think in the absence of contrary evidence, they are to be regarded as some proof.[1]

1 *Cheatle v Considine* [1965] SASR 281, 282.

Presumptions and mechanical instruments

6.5 The presumption that scientific instruments work properly has a long history.[1] For instance, scales benefit from the presumption.[2] Timing devices also take advantage of the presumption. In *Plancq v Marks*,[3] in an appeal against conviction for driving a motor car in excess of the speed limit of 20mph, the evidence of the police officer was challenged. The stop watch used by the police officer was produced in court. The appeal focused on the ground that the police officer gave opinion evidence as to the speed of the vehicle. This appeal was dismissed on the basis that the police officer was merely reading out the reading from the stop watch, which did not constitute the giving of opinion evidence; the real issue was whether the police officer was telling the truth.

1 R P Groom-Johnson and G F I Bridgman (eds), *A Treatise on the Law of Evidence* (12th edn, Sweet and Maxwell 1931), in which the working accuracy of certain scientific instruments were recognized in the absence of evidence to the contrary, such as watches, clocks, thermometers, aneriods and anemometers, amongst other 'ingenious contrivances', 167.
2 *Giles v Dodds* [1947] VLR 465, [1947] ArgusLawRp 53; (1947) 53 Argus LR 584.
3 (1906) 94 LT NS 577.

6.6 Arguments that a watch used to prove that the defendant was speeding ought to be tested have been ignored,[1] as in the case of *Gorham v Brice*.[2] The Lord Chief Justice dismissed the appeal against conviction for driving a motor car in excess of the speed limit of 12 mph without considering the point. In comparison, the members of the Divisional Court in *Melhuish v Morris*[3] allowed an appeal against speeding because the speedometer of the police vehicle had not been tested for accuracy.[4] The court in *Nicholas v Penny*[5] subsequently overturned this decision. Lord Goddard CJ commented as follows:

> The question in the present case is whether, if evidence is given that a mechanical device, such as a watch or speedometer – and I cannot see any difference in principle between a watch and a speedometer – recorded a particular time or a particular speed, which is the purpose of that instrument to record, that can by itself be prima facie evidence, on which the court can act, of that time or speed.[6]

1 In communication with the author, Professor Strigini points out that from an engineering point of view, testing that a watch is accurate enough now (which usually implies that it was until now, unless it has been repaired) is inexpensive enough that not doing it seems a dereliction of duty.
2 (1902) 18 TLR 424.
3 [1938] 4 All E R 98; see also 'Evidence in speed limit cases' (1937) 1 Journal of Criminal Law 181.
4 Evidence that the accused did not exhibit the usual signs of being intoxicated can indicate that a machine is not working properly: *DPP v Spurrier* [2000] RTR 60; police officers can conduct physical tests to ensure a speedometer is working accurately, for which see *Pervez v Procurator* [2000] ScotHC 111.
5 [1950] 2 All ER 89, DC; *Penny v Nicholas* [1950] 2 KB 466; 66 Law Quarterly Review (1950) 264, 441; in the South Australian case of *Peterson v Holmes* [1927] SASR 419, Piper J asked, at 421, 'If [the word 'It' is in the report, but must be a mistake] the speedometer be tested by stop-watches and measured distances, what about the accuracy of the watches and the chain measure?' ('Proof of excessive speed' (1950) 14 Journal of Criminal Law 360).
6 [1950] 2 KB 466, 473.

6.7 The judge went on to suggest that because the defendant was accused of exceeding the speed limit by 10 mph, it 'would be a considerable error in the speedometer if it were as much out as that'.[1] Such a comment was not intended, it is suggested, to create a presumption that such devices are reliable, especially as Lord Goddard CJ commented that 'the justices need never accept any evidence if they do not believe it, or feel that for some reason they cannot accept it'.[2] A similar issue arose in the case of *H. Gould and Company Limited v Cameron*,[3] where the tyres of a heavy motor-vehicle were tested in July and found to be over the legal limit. The instrument used to test the tyre pressure had been tested in March of the previous year, and in August in the year following the reading. The defence argued that the instrument might have developed an error after testing in March. It was known and accepted that, at certain pressures, the device would be in error of 1lb over a range of tests between 70lb and 100lb. This error had been taken into account in this case. Northcroft J said:

> In a case such as this, where of necessity, a mechanical device must be used to ascertain the pressure within the tyres, it is sufficient, I think, to show that the instrument is used correctly, and that, from its nature and history, it may reasonably be relied upon by the Court. The history of this instrument and the description of its use satisfies me that the learned Magistrate was justified in accepting it, as I do, as a reliable test on this occasion.[4]

1 [1950] 2 KB 466, 473.
2 [1950] 2 KB 466, 742. In *R v Amyot* (1968) 2 O.R. 626, Clare Co.Ct.J accepted the use of a stop watch

to measure the time a vehicle took to travel between marked points on a highway, where the police officer had personally checked the distance between the markings using a cyclometer and made the observations with the stop watch in an aircraft.

3 [1951] NZLR 314.

4 [1951] NZLR 314, 316 (40 – 45).

6.8 The observations by Shadbolt DCJ in the New South Wales case of *Re Appeal of White*[1] put the matter into perspective when hearing an appeal for exceeding the speed limit, where he noted:

> Courts have been generally loath to be wearied in seeking proof of some absolute measure or requiring it in cases such as this. It is not possible for every child to check his wooden ruler with the standard metre in Canberra nor every grocer his scales with the standard gram. Most of us accept the ruler's accuracy and the weight of the grocer's scales.

1 (1987) 9 NSWLR 427, 430.

6.9 Therefore, it does not follow that every measuring device is accurate.

Judicial formulations of the presumption that mechanical instruments are in order when used

Judicial notice

6.10 There are a number of reasons for the doctrine of judicial notice:[1] to expedite the hearing of a case where obvious facts do not need proving; to promote uniformity in judicial decision making, and to prevent the possibility of a decision which is demonstrably erroneous or false.[2] Brett JA summed up the concept in *R v Aspinall*: 'Judges are entitled and bound to take judicial notice of that which is the common knowledge of the great majority of mankind and of the greater majority of men of business.'[3] In the High Court of Australia,[4] Isaacs J emphasized the guiding principle of the doctrine:

> The only guiding principle—apart from Statute—as to judicial notice which emerges from the various recorded cases, appears to be that wherever a fact is so generally known that every ordinary person may be reasonably presumed to be aware of it, the Court 'notices' it, either *simpliciter* if it is at once satisfied of the fact without more, or after such information or investigation as it considers reliable and necessary in order to eliminate any reasonable doubt.
>
> The basic essential is that the fact is to be of a class that is so generally known as to give rise to the presumption that all persons are aware of it.[5]

1 See Law Commission New Zealand, 'Evidence law: documentary evidence and judicial notice. A discussion paper' (Preliminary Paper No 22, 1994), ch. IX for a more nuanced consideration of the topic; Hodge M Malek (ed), *Phipson on Evidence* (18th edn, Sweet & Maxwell 2013) ch. 3.

2 Tapper, *Cross and Tapper on Evidence* 84; for examples, see 77 and Christopher Allen, 'Case Comment: Judicial notice extended' (1998) 2 E & P 37, 39; David M Paciocco, 'Proof and progress: coping with the law of evidence in a technological age' (2013) 11 Canadian Journal of Law and Technology 181, 188–9; Evidence (Interim) [1985] ALRC 26, [969]; Law Commission New Zealand, 'Evidence law: documentary evidence and judicial notice' [259].

3 (1876) 3 QBD 48, 61 – 62.

4 *Holland v Jones* 23 CLR 149 (1917), [1917] VLR 392, 23 ALR 165, 1917 WL 15976, [1917] HCA 26.

5 23 CLR 149 (1917), 153.

6.11 Lord Summer considered the practical approach in *Commonwealth Shipping Representative v P. & O. Branch Service*:

> My Lords, to require that a judge should affect a cloistered aloofness from facts that every other man in Court is fully aware of, and should insist on having proof on oath of what, as a man of the world, he knows already better than any witness can tell him, is a rule that may easily become pedantic and futile.[1]

1 [1923] AC 191, 211.

6.12 The doctrine of judicial notice is restricted to very clear knowledge,[1] and it can be more severe in its effect than a presumption, as noted by Susan G. Drummond:

> It is a manoeuvre that forecloses further evidence. The judge operates, in this case, as a virtually unlimited authority with limitations imposed only from within the legal hierarchy. Judicial notice can only be contested on appeal and invalidated if it can be demonstrated that the criteria for the application of judicial notice were not present (the fact was not notorious, the sources to establish the fact were not indisputable ...). As judicially noticed matters operate in the domain of fact, not law, they have no precedential value.[2]

1 For discussions on the confusing treatment of this doctrine, see G D Nokes, 'The limits of judicial notice' (1958) 74 Law Quarterly Review 59 and Susan G Drummond, 'Judicial notice: the very texture of legal reasoning' (2000) 15 Canadian Journal of Law and Society 1.
2 Drummond, 'Judicial notice' 4.

6.13 Given that it appears as if this doctrine has been extended to electronic evidence in Canada, this observation by Drummond illustrates the importance of ensuring judges more fully understand the nature of the world in which they now live. Thorson JA discussed judicial notice in *R. v Potts* before the Ontario Supreme Court, Court of Appeal:[1]

> Judicial notice, it has been said, is the acceptance by a court or judicial tribunal, without the requirement of proof, of the truth of a particular fact or state of affairs that is of such general or common knowledge in the community that proof of it can be dispensed with.
>
> …
>
> Thus it has been held that, generally speaking, a court may properly take judicial notice of any fact or matter which is so generally known and accepted that it cannot reasonably be questioned, or any fact or matter which can readily be determined or verified by resort to sources whose accuracy cannot reasonably be questioned.

1 1982 CarswellOnt 56, [1982] OJ No. 3207, 134 DLR (3d) 227, 14 MVR 72, 26 CR (3d) 252, 36 OR (2d) 195, 66 CCC (2d) 219, 7 WCB 236, at [15].

6.14 In *R. v Find*,[1] before the Supreme Court of Canada, McLachlin CJC directed that the threshold for judicial notice is strict:

> Judicial notice dispenses with the need for proof of facts that are clearly uncontroversial or beyond reasonable dispute. Facts judicially noticed are not proved by evidence under oath. Nor are they tested by cross-examination. Therefore, the threshold for judicial notice is strict: a court may properly take judicial notice of facts that are either: (1) so notorious or generally accepted as not to be the subject of debate among reasonable persons; or (2) capable of immediate and accurate demonstration by resort to readily accessible sources of

indisputable accuracy.[2]

1 2001 CarswellOnt 1702, 2001 CarswellOnt 1703, 2001 SCC 32, [2001] 1 SCR 863, [2001] SCJ No. 34, 146 OAC 236, 154 CCC (3d) 97, 199 DLR (4th) 193, 269 NR 149, 42 CR (5th) 1, 49 WCB (2d) 595, 82 CRR (2d) 247, J.E. 2001-1099, REJB 2001-24178.
2 At [48].

6.15 The concept of 'notorious' is considered in *Phipson*:

> the concept covers matters being so notorious or clearly established or susceptible of demonstration by reference to a readily obtainable and authoritative source that evidence of their existence is unnecessary. Some facts are so notorious or so well established to the knowledge of the court that they may be accepted without further enquiry.[1]

1 Malek (ed.), *Phipson on Evidence* para 3:02.

6.16 The judge can conduct their own research, and the United States Court of Appeals, Ninth Circuit reached conclusions regarding automatic programs in this way, as in *U.S. v Lizarraga-Tirado*, where Kozinski, CJ said:

> Because there was no evidence at trial as to how the tack and its label were put on the satellite image, we must determine, if we can, whether the tack was computer generated or placed manually. Fortunately, we can take judicial notice of the fact that the tack was automatically generated by the Google Earth program. By looking to 'sources whose accuracy cannot reasonably be questioned'—here, the program—we can 'accurately and readily determine[ ]' that the tack was placed automatically. *See* Fed.R.Evid. 201(b). Specifically, we can access Google Earth and type in the GPS coordinates, and have done so, which results in an identical tack to the one shown on the satellite image admitted at trial.[1]

1 789 F.3d 1107 (9th Cir. 2015), 1109.

6.17 In justifying judicial notice, David M. Paciocco comments that 'If a court could not rely on a notorious and incontrovertible material fact because it had not been proved, verdicts would not conform to reality. The repute of the administration of justice would be harmed'.[1] Paciocco went on to illustrate his argument with the following example of how a brake on a motor vehicle operates:

> For example when someone describes putting the brakes on in a car no-one offers expert testimony that the function of brakes is to slow or stop vehicles, that brakes are typically controlled by foot-pedals that are depressed in order to slow or stop the vehicle, or that brakes are depressed gently to come to a gradual stop and aggressively for an emergency stop.[2]

1 Paciocco, 'Proof and progress' 188–9.
2 Paciocco, 'Proof and progress' 189

6.18 But there is a distinction between the purpose of a brake on a motor vehicle and how the braking system operates. In the example above, Paciocco made assumptions about how braking systems work and failed to understand the nature of the technology. Most braking systems in motor vehicles are controlled by a mix of electronic systems and software code (a fact so notorious that no citation ought to be required[1]). It is more accurate, using a high level functional description of the brake system, to explain the braking technology in vehicles as involving the use of brakes primarily under the control of electronics or software code. The failsafe fallback strategy is that if the

electronics or software code fails, the system reverts to a standard hydraulic brake system. It does not necessary follow that the function is always performed correctly or as normally expected in the situation where the action is mediated by electronic systems. For instance, antilock braking systems (ABS), electronic stability control (ESC) and traction control are predicated on interactions between the engine torque output and brake control on individual wheels. This means that there is a possible difference between the fact that a braking event took place, and whether or not a braking event was requested, and vice versa.[2] This example is far from the strict application of the doctrine as noted in the Supreme Court of Canada by McLachlin CJC. If judicial notice is extended to such an extent, then the question of whether justice is served by this doctrine must be carefully scrutinized.

1 Notwithstanding it is notorious that anti-lock brake systems are partly controlled by software code and electronic systems, the reader can obtain more information from the *Society of Automotive Engineers International*; the open access journal *Intelligent Control and Automation*, and *IEEE Transactions on Vehicular Technology*.

2 I owe this point to Dr Michael Ellims; see also the following, in which it is demonstrated that braking systems can be controlled by hacking into the motor vehicle computer system: C. Valasek and C. Miller, 'Adventures in automotive networks and control units' (Technical White Paper, 2014) <www.ioactive.com/pdfs/IOActive_Adventures_in_Automotive_Networks_and_Control_Units.pdf>; Charlie Miller and Chris Valasek, 'Remote exploitation of an unaltered passenger vehicle' (2015) <http://illmatics.com/Remote%20Car%20Hacking.pdf>; Roderick Currie, 'Developments in car hacking' (SANS Institute, 2015) <www.sans.org/reading-room/whitepapers/internet/developments-car-hacking-36607>.

A 'notorious' class

6.19 In the Victoria case of *Crawley v Laidlaw*,[1] Lowe J considered the basis upon which a presumption might apply – in this case regarding a scientific instrument:

> I do not question that such a presumption is frequently and (in general) tacitly acted on by our Courts; but in my opinion it must appear from evidence before the Court, or from something which stands in place of evidence, *e.g.*, judicial notice, that the instrument in question is a scientific instrument, before the presumption applies.

1 (1930) VLR 370, 374.

6.20 The prosecution sought to adduce evidence from two weighing machines called 'loadometers' to prove a motor truck was carrying a greater weight than that allowed by the regulations. The Police Magistrate who heard the case had dismissed it on the basis that there was no evidence to demonstrate the correctness of the instruments. On appeal, Lowe J concurred, holding that there was no evidence that the devices were scientific instruments, and there was no foundation for a presumption that the instrument worked properly. Emphasising the need to establish a foundation for the presumption, Lowe J observed:

> I do not doubt that in appropriate cases the Court will use its 'general information and … knowledge of the common affairs of life which men of ordinary intelligence possess' – *Phipson on Evidence* (6th ed.), p. 19 – and that of the nature of most, if not all, of the instruments mentioned in the paragraph cited from *Taylor*[1] the Court would require no evidence in order to raise the presumption relied on. I think, too, that the Court may, if it thinks it desirable, refer to appropriate standard works of reference in order to inform itself of matters of the kind mentioned of, which it may personally be unaware. But if, after such reference, the Court is still

> ignorant of the nature of the instrument in question, no help can be got from the presumption relied on. Apparently the learned magistrate did not know, and I myself do not know, what a loadometer is. I may guess, from the derivation of the name what the instrument is, but my guess is not evidence.[2]

1 '*Taylor on Evidence* (10th ed.), sec. 183, where the learned author says: "The working accuracy of scientific instruments is also presumed. For example, in the absence of evidence to the contrary, a jury would be advised to rely on the correctness of a watch or clock, which had been consulted to fix the time when a certain event happened; a thermometer would be regarded as a sufficiently safe indication of the heat of any liquid in which it had been immersed; a pedometer would afford prima facie evidence of the distance between two places which had been traversed by the wearer; and similar prima facie credit would be given to aneroids, anemometers, and other scientific instruments; and blood stains are every day detected by means of known chemical tests."'(1930) VLR 370, 373–4. The quote in the footnote uses the term 'correctness'; others more correctly seem to refer to 'sufficient accuracy'. A measurement instrument for a continuous quantity has a degree of accuracy (how close the reading is to the real value) and a degree of precision (how tightly spaced the points are on its scale), but its reading will not usually be exactly 'correct'. This may have a bearing on how digital devices are seen. A tiny damage to the mechanical mechanism of a scale might cause it to be slightly off the right reading of weight, but a tiny mistake in software may change the response to some specific inputs substantially. I owe this to Professor Strigini.
2 (1930) VLR 370, 374.

6.21 Herring CJ made comments similar to Lowe J's in the Victoria case of *Porter v Koladzeij*.[1] This case involved the review of the refusal of a Stipendiary Magistrate to admit evidence of an analogue device to measure the amount of alcohol in a sample of breath. The judge observed that certain instruments of a scientific or technical nature fell into a 'notorious' class that by general experience are known to be trustworthy. He placed a speedometer into this class. However, the evidence from the device to measure breath alcohol was rejected because it was not a standard device, and because the evidence given by the witness regarding the device was not adequate. The judge said that once breath analysis devices were used more often, they would become standard, and then judicial notice would be taken of their existence as scientific or technical instruments,[3] although it was necessary to present relevant evidence to the court:

> Where, however, the instrument in question does not fall within the notorious class, then his Honour made it clear that evidence must be given to establish that it is a scientific or technical instrument of such a kind, as may be expected to be trustworthy, before the presumption can be relied upon.[4]

1 (1962) VR 75.
2 Falling back on 'general experience' is dubious, because few people check the correctness of the instruments they might use. People routinely use imprecise instruments such as house thermometers, speedometers and seldom have occasions for questioning the readings – but by relying on the reading does not make the reading accurate.
3 The Supreme Court in South Australia refused to take judicial notice of the accuracy of the breathalyser in 2012: *Police v Bleeze* [2012] SASCF 54, [88] and [89].
4 (1962) VR 75, 78.

6.22 The failure to obtain such evidence can lead, as described by Thomas E. Workman, to scenarios such as that described below:

> In Florida, one citizen was tested 13 times on one machine, by one officer, in one hour. These instances occur because in some situations, a machine that registers an error or multiple errors may finally produce a value that has the appearance of being a valid test. The Courts are usually unaware of the history of failures on the machine, and believe that the result is legitimate, when in fact may not be.[1]

1 Thomas E Workman, Jr, 'Massachusetts breath testing for alcohol: a computer science perspective' (2008) 8 Journal of High Technology Law 209, 217.

6.23 However, it is not necessary to rely on a presumption that an instrument is accurate or reliable in lieu of other evidence that the data produced by the instrument is accurate.[1] For instance, a satellite navigation system was the subject of discussion in *Chiou Yaou Fa v Thomas Morris*[2] before the Supreme Court of the Northern Territory of Australia. In this case, the commander of the vessel established his position by way of the satellite navigation system, radar and sextant. The court accepted the evidence that a variety of methods were used to establish the position at sea, including the expertise of qualified navigators. Even though the court heard their testimony as to the accuracy of the satellite navigation system, it concluded that it was not necessary to rely upon the satellite navigation system as being in the 'notorious' class, and accepted the radar and sextant evidence in its place.[3]

1 In *R. v Ranger* 2010 CarswellOnt 8572, 2010 ONCA 759, [2010] OJ No. 4840, 91 WCB (2d) 271, the Ontario Court of Appeal held at [16] 'it is now notorious that cell phone users engaged in a cell phone call and travelling from point A to point B will find their cell phone signal passes from one cell phone tower to another at different locations along the route from point A to point B', which led the court to consider that the trial judge did not err 'in taking judicial notice that a particular cell phone was in a general location based on the tower that received the signal and that the path along which the cell phone was moving could be determined by reference to the cell phone towers that received the signal transmission in respect of particular calls'.

2 (1987) 46 NTR 1.

3 United States of America: *St. Martin v Mobil Exploration & Producing U.S. Inc.*, 224 F.3d 402 (5th Cir. 2000) 31 Envtl. L. Rep. 20, 01155 Fed. R. Evid. Serv. 270 (aerial photography); *Connecticut v Wright*, 58 Conn.App. 136, 752 A.2d 1147 (Conn.App. 2000) (computer generated engineering map); *Wetsel-Oviatti Lumber Co. Inc., v United States*, 40 Fed.Cl. 557 (1998) (aerial photography); *United States v Kilgus*, 571 F.2d 508 (9th Cir. 1978) (infrared rays); *Pittson Co. v Allianz Insurance Co.*, 905 F.Supp. 1279 (D.N.J. 1995) rev'd in part on other grounds, 124 F.3d 508 (3d Cir. 1997) (aerial photography); *Ponca Tribe of Indians of Oklahoma v Continental Carbon Co.*, 2008 WL 7211981 (digital orthophoto); *Gasser v United States*, 14 Cl.Ct. 476 (1988) (aerial and satellite photographs); I & M Rail Link v Northstar Navigation, 21 F.Supp. 849 (N.D.Ill. 1998) (satellite photography); *Wojciechowicz v United States*, 576 F.Supp.2d 214 (D.Puerto Rico 2008) (satellite photography); *Lisker v Knowles*, 651 F.Supp.2d 1097 (C.D.Cal. 2009) (satellite photography); *United States v Fullwood*, 342 F.3d 409 (5th Cir. 2003) (satellite photography); *Fry v King*, 192 Ohio App.3d 692, 950 N.E.2d 229 (Ohio App. 2 Dist. 2011), 2011 WL 766583 (satellite photography); *State v Reed*, 2009 WL 2991548 (Google Earth evidence rejected); *State of New Jersey in the Interests of J. B. A Minor*, 2010 WL 3836755 (Google Earth evidence admitted); *Swayden v Ricke*, 242 P.3d 1281 (2010), 2010 WL 4977158 (Google Earth images and photographs from 'trail cameras'); *Banks v U.S.*, 94 Fed.Cl. 68 (2010) (satellite photography).

Common knowledge

6.24 Another justification for accepting that a mechanical instrument is in order when it is used is the assertion that it is a type of instrument that is commonly held to be – more often than not – in 'working order'. In a case before the full court of the Supreme Court of Western Australia, *Zappia v Webb*,[1] the question was whether an amphometer, used to determine the speed of a vehicle, could be considered an accepted scientific instrument. Jackson CJ discussed this as follows:

> It is, however, common knowledge that amphometers have been widely used in this State for a number of years for the purpose of checking the speed of motor vehicles. As one drives through the country, it is common-place to see large notices by the side of the road warning motorists that amphometers are used

> in the district, and it is not at all uncommon to see a traffic inspector by the side of the road with his amphometer equipment set up. It is also, I believe, generally accepted in the community that an amphometer correctly set up and operated will give a reliable reading of speed, not necessarily precise, but sufficiently accurate for its purpose. There has not been, so far as I am aware, any general complaint about the use or efficiency of these machines, and there must be hundreds of speeding convictions each year resulting from their use.
>
> It seems to me, therefore, that an amphometer is now a well known and accepted speed checking device and that judicial notice should be taken in this State of its use and effectiveness, in general terms.[2]

1 (1974) WAR 15; (1973) 29 LGRA 438.
2 (1973) 29 LGRA 438, 440 – 441.

6.25 The Chief Justice referred to the 'common knowledge' of the use of amphometers without referring to any evidence to demonstrate that they were reliable. He also asserted that somehow it was generally accepted that the device would give a reliable reading of speed, and concluded that because he was not aware of any complaints about the devices, they were therefore to be considered an accepted speed checking device.

6.26 In *Castle v Cross*, the prosecution relied on the presumption that mechanical instruments were in order when they were used. In the judgment, Stephen Brown LJ cited a passage from *Cross on Evidence* (1979)[2] regarding this presumption:

> A presumption which serves the same purpose of saving the time and expense of calling evidence as that served by the maxim omnia praesumuntur rite esse acta is the presumption that mechanical instruments were in order when they were used. In the absence of evidence to the contrary, the courts will presume that stopwatches and speedometers and traffic lights were in order at the material time; *but the instrument must be one of a kind which it is common knowledge that they are more often than not in working order*.[1] (emphasis added)

1 [1984] 1 WLR 1372; [1985] 1 All ER 87, QBD.
2 Page 47 of the fifth edition.
3 [1984] 1 WLR 1372, 1376H – 1377A.

6.27 The Latin tag 'omnia praesumuntur rite esse acta' means 'all acts are presumed to have been done rightly and regularly' or 'all things are presumed to have been done regularly and with due formality until the contrary is proved'. Such a presumption cannot operate in a vacuum, as indicated by Stephen Brown LJ's preference for the above formulation in *Cross on Evidence* which requires the basic fact – proof that the instrument be one of a kind which is common knowledge that they are more often than not in working order – to be established before the presumption could operate, as opposed to the same formulation of the presumption in *Phipson on Evidence*, which did not adopt the basic fact.[1]

1 [1984] 1 WLR 1372, 1377.

6.28 In this case, counsel for the Crown put forward the case that the device in question, a Lion Intoximeter 3000, was a sophisticated machine that depended in part on software code, but this did not set it in a different class from other sophisticated mechanical devices and instruments. The presumption stood unchallenged because

the defence 'argued forcefully that the potential for computer error renders the consideration of evidence stemming from a computer particularly sensitive and places it into a separate class in relation to its admissibility'.[1] It is unclear from the judgment of Stephen Brown LJ whether his Lordship relied on the presumption in admitting the print-out from the Lion Intoximeter 3000, as the central issue in this case appears to be the admissibility of the print-out as real evidence.

1 [1984] 1 WLR 1372, 1379D.

6.29 *Anderton v Waring* also concerned the reading from a Lion Intoximeter 3000. In giving the judgment of the court, May LJ stated that the 'Intoximeter ought to have been assumed by the justices to have been in good working order unless the contrary was proved'.[2] When counsel for the prosecution cited from the fourth edition of *Cross on Evidence*:[3] 'In the absence of evidence to the contrary, the courts will presume that mechanical instruments) were in order at the material time',[4] counsel omitted to cite the basic fact that '… the instrument must be one of a kind as to which it is common knowledge that they are more often than not in working order'.[5] This has to be a misapplication of the presumption because a presumption cannot operate in a vacuum without the basic fact or facts.

1 [1986] RTR 74.
2 [1986] RTR 74, 80F.
3 Page 47.
4 [1986] RTR 74, 79E.
5 (6th edn, 1985), 28; Professor Tapper mentioned this omission in Colin Tapper, 'Reform of the law of evidence in relation to the output from computers' (1995) 3 Intl J L & Info Tech 79, 89.

Properly constructed

6.30 A more recent presumption has been articulated by Kerr LCJ, as he then was, when he rejected the suggestion that the machine in question ought to be commonly known to be – more often than not – in working order. In *Public Prosecution Service v McGowan*, Kerr LCJ said:

> In so far as the passage from *Cross and Tapper* suggests that for the presumption to operate *it will always be necessary that the machine was commonly known to be more often than not in working order, we would not accept it*. We consider that the presumption must be that machines such as a cash register are operating properly and in working order in the absence of evidence to the contrary. The presumption of the correct operation of equipment and proper setting is a common law presumption recognised by article 33(2) [Criminal Justice (Evidence) (Northern Ireland) Order 2004]. In the modern world the presumption of equipment being properly constructed and operating correctly must be strong.[1]

1 [2008] NICA 13, [2009] NI 1, [20].

6.31 Kerr LCJ's deviation from the formulation of the presumption, which requires proof of the basic fact, appears to be unwarranted. Furthermore, Kerr LCJ's formulation of the presumption without the basic fact leads to the extraordinarily broad assumption that all devices and machines are operating properly and in working order, an assumption for which his Lordship did not cite any scientific evidence in support. In particular, there was nothing in the judgment to indicate what he knew by 'equipment',

or how the equipment was 'properly constructed'. Nor did he provide any evidence as to what he meant by 'operating correctly' or 'proper setting'.

Evidential foundations of the presumption

6.32 It is suggested that the correct articulation of the mechanical instruments presumption is that as indicated above in *Crawley v Laidlaw*[1] and *Porter v Koladzeij*,[2] which is:

> For a mechanical instrument (including stand-alone computers, computer-like devices and digital systems) to benefit from the evidential presumption that it was in working order at the material time, it is necessary for the partly seeking to benefit from the presumption to adduce evidence of how the instrument in question works, and to include reference to relevant scientific papers and texts that support such an assertion.

1 (1930) VLR 370.
2 (1962) VR 75.

6.33 This formulation is consistent with *Crawley v Laidlaw*[1] and *Porter v Koladzeij*[2] in that if the presumption is to be recognized, it is necessary for the proponent to provide sufficient evidence – the basic fact – to merit the introduction of such a presumption. It this respect, it is pertinent to note the observation by Lord Griffiths in *Cracknell v Willis*[3] that '"trial by machine" is an entirely novel concept and should be introduced with a degree of caution'. He went on to indicate that it would be unthinkable that somebody should be convicted by a machine that is not 'reliable', although he did not make it clear what he meant by 'reliable'. The basic fact in the maxim omnia praesumuntur provides a simple (but not infallible) yardstick to assess if the machine is 'reliable': it must 'be one of a class of machines which it is common knowledge that they are more often than not in working order'.

1 (1930) VLR 370.
2 (1962) VR 75.
3 [1988] 1 AC 450 at 459, [1987] 3 All ER 801 at 806, HL; work had already been undertaken before 1988: T R H Sizer and A Kelman (eds), *Computer generated output as admissible evidence in civil and criminal cases* (Heyden & Son on behalf of the British Computer Society 1982); Alistair Kelman and Richard Sizer, *The Computer in Court* (Gower 1982).

6.34 Conversely, in *DPP v McKeown; DPP v Jones*[1] Lord Hoffmann voiced the opinion in 1997 that 'It is notorious that one needs no expertise in electronics to be able to know whether a computer is working properly'. This comment, akin to the 'aura of infallibility',[3] is an extreme view that is contradicted by the technical evidence, and does not bear a great deal of scrutiny. The observation by Lloyd LJ in *R v Governor Ex p Osman (No 1), sub nom Osman (No 1), Re*[4] is of a similar nature:

> Where a lengthy computer printout contains no internal evidence of malfunction, and is retained, e.g. by a bank or a stockbroker as part of its records, it may be legitimate to infer that the computer which made the record was functioning correctly.

1 [1997] 1 All ER 737, [1997] 1WLR 295, HL; also note the comment by Harvey J in the New Zealand case of *R v Good* [2005] DCR 804 at 65 'that computers are not recently invented devices, are in wide use and are fundamentally reliable'.

2 [1997] 1 All ER 737, 743b.
3 D W Elliott, 'Mechanical aids to evidence' [1958] Crim LR 5, 7.
4 [1989] 3 All ER 701, [1990] 1 WLR 277, (DC), 306H.

6.35 The judge did not indicate what evidence was before him to demonstrate that there was no 'internal evidence of malfunction', and just because the bank or a stockbroker relied on computer data as part of its records – or, as George L. Paul puts it, '[j]ust because businesses rely on faulty computer programs does not necessarily mean that courts should follow suit'. Indeed, Professor Seng observed that such comments made by judges are '… extravagant judicial statements … [that] are incomplete and are actually misleading because accurate computer output depends not just on the proper operation of computers, but also proper human use (or abuse) of computers'.[2]

1 George L Paul, *Foundations of Digital Evidence* (American Bar Association 2008) 129; *Gordon v Thorpe* [1986] RTR 358 where two experts gave evidence of the accuracy or otherwise of a Lion Intoximeter 3000.
2 Daniel K B Seng, 'Computer output as evidence' (1997) Sing JLS 130, 167.

6.36 The whole idea of 'instrument in working order' relies on the presumption that transitions between 'being in working order' and 'not being in working order' are reasonably rare. In other words, that the instrument cannot capriciously alternate between giving correct readings and incorrect readings, with arbitrary lengths of the sequences of correct and of incorrect readings. These arbitrary sequences may happen with software. Although there is generally a reason for these sequences – something in the exact values and timings of the sequences of inputs determines which outputs will be correct and which ones will be wrong, given the defects in the software, identifying the law that governs them and the software defects causing it may be impossibly time-confusing even for well-equipped experts.

How judges assess the evidence

6.37 When discussing the admission of evidence from devices controlled by software code, judges do not distinguish between a single, highly specialist device that is self-contained, and a linked network containing any number of devices each independently operating on its own set of software code. As noted above, when considering cases dealing with specialized devices such as breath testing machines and blood testing machines, judges have used nebulous terms in the absence of scientific analysis, using such terms as such as 'notoriety', 'common knowledge' and 'properly constructed'. There is little evidence to demonstrate that proper evidential foundations have been adduced to permit such presumptions to be admitted. In this regard, it is useful to consider, although not exclusively, the case law in Australia, where these devices have been subjected to greater judicial analysis.

6.38 The Southern Australian case of *Mehesz v Redman*[1] was a case that concerned the method of analysing a blood sample. At trial, the Special Magistrate categorized the blood sample testing device as a scientific instrument with the presumption that it was in the category of a 'notorious' instrument whose accuracy is presumed. On appeal, Zelling J rejected this on the basis that the device was not a mere calculator, although it interpreted the results because of the software program, because there was

no evidence to demonstrate that the machine was accurate or reliable. The appellant was tried a second time, convicted again, and appealed to the Supreme Court once more. This appeal was referred to the full court.[2] The main argument of counsel for the appellant related to the evidence tendered by the prosecution regarding the analysis of a blood sample, in that the evidence relied on the use of two instruments (a gas chromatograph and the 'Auto-lab system 4B' data analyser) whose accuracy had not been established. King CJ rejected the submission that the Auto-lab was an instrument that could not be relied upon because there was no evidence as to the 'correctness' of the software program. He said:

> The courts do not require such evidence. If the instrument is so well known that its accuracy may be assumed as a matter of common experience, the Court is entitled to presume its accuracy without evidence.[3]

1 (1979) 21 SASR 569.
2 *Mehesz v Redman (no 2)* (1980) 26 SASR 244.
3 (1980) 26 SASR 244, 247.

6.39 Proof of the accuracy of a particular instrument will 'ordinarily be proved by those who use and test it', and the results obtained are acceptable in evidence 'provided that the expert witness has himself formed an opinion that the methods used are apt to produce the correct result'.[1] Notwithstanding the inability of the operator of a machine controlled by software code to demonstrate the accuracy or otherwise of the code that he does not control and has no ability to alter, this proviso is important. (White J also made a similar point.[2]) This means that the operator of such a machine ought to be able to assess when the machine produces results that are not expected, even if the operator is not able to establish why the results produced are wrong. If such a machine produces results that are not anticipated, the operator is put on notice that the machine (and the software code) might not be reliable. In such circumstances, it will be necessary to have the machine tested before being relied upon for future analysis.

1 (1980) 26 SASR 244, King CJ at 248.
2 (1980) 26 SASR 244, 254.

6.40 Dealing with the submission that the prosecution failed to provide proper foundations for the Auto-lab analyser, White J set out the conditions that must be fulfilled before evidence will be admitted regarding the measurements of scientific instruments:

> 1. If the instrument falls within the class of instrument known as notorious scientific instruments, the court will take judicial notice of its capacity for accuracy, so that the operator merely proves that he handled it properly on the particular occasion.
>
> 2. If the instrument is *not* a notorious scientific instrument, its accuracy can be established by evidence: (a) that the instrument is within a class of instrument generally accepted by experts as accurate for its particular purpose; (b) that the instrument, if handled properly, does produce accurate results: ((a) and (b) must be established by expert testimony, that is, by experts with sufficient knowledge of that kind of instrument; and upon proof of (a) and (b), a latent *presumption of accuracy* arises which allows the court to infer accuracy on the particular occasion if it is proved) – (c) that the particular instrument was handled properly and read accurately by the operator on the particular occasion; ((c) can be established by

> a trained competent person familiar with the operation of the instrument, not necessarily the type of expert who proves (a) and (b)).
>
> 3. Where the actual accuracy of the measurement can be inferred from all of the proved circumstances, it is not necessary to rely upon the presumption arising from (a) and (b), proof of which is superfluous.[1]

1 (1980) 26 SASR 244, 251 – 252.

6.41 At the second trial, the prosecution called evidence from Professor Northcote, Chairman of the School of Mathematics and Computers at the Institute of Technology in South Australia, and an expert in mathematics, physics and computers. He gave evidence about the workings of the Auto-lab from his reading of the manufacturer's manual and his understanding of the content of the manual. He was not able to read the software code, because the manufacturer had sealed the program against inspection, tampering and modification. Although Professor Northcote was not an expert in relation to the Auto-lab, the members of the Court of Appeal in the Supreme Court were of the option that both Professor Northcote and Mr Vozzo, who gave evidence at both trials, were sufficiently qualified to give evidence, even though neither witness had access to, nor any knowledge of, the software code. The Chief Justice also stated that 'It is sufficient that the expert who uses it is able to say that it is an instrument which is accepted and used by competent persons as a reliable aid to the carrying out of the scientific procedures in question and that he so regards it.'[1] He also prayed in aid the observations of *Wigmore on Evidence* to support this comment:

> (2) Scientific instruments, formulas, etc. The use of *scientific instruments, apparatus, formulas, and calculating-tables,* involves to some extent a dependence on the statements of other persons, even of anonymous observers. Yet it is not feasible for the professional man to test every instrument himself; furthermore he finds that practically the standard methods are sufficiently to be trusted. Thus, the use of a vacuum-ray machine may give correct knowledge, though the user may neither have seen the object with his own eyes nor have made the calculations and adjustments on which the machine's trustworthiness depends. The adequacy of knowledge thus gained is recognized for a variety of standard instruments.[3]

1 (1980) 26 SASR 244, 247.
2 (3rd ed), Volume 2, paragraph 665a.
3 (1980) 26 SASR 244, 247.

6.42 In this case, the court emphasized that there was evidence other than the trustworthiness of the software code that enabled the evidence from the machine to be admitted as being accurate. White J set out the following analysis of the problem:

> The only defect in the expert evidence of Dr. Northcote and Mr. Vozzo, if defect it be, was their lack of direct knowledge of the internal operations of the sealed instrument. They relied upon what the manufacturer said about its operation. The extreme position would be that only the expert actually supervising the manufacture of the instrument in the United States of America could prove (a) and (b). I do not think that the rules relating to expert evidence encourage that kind of extreme position. Quite apart from questions of expense and delay in the administration of justice, the Court is entitled to rely upon evidence of measurements made by instruments which reputable scientists accept as accurate, whether those scientists have direct knowledge of the reasons for the instrument's accuracy or not, provided they have knowledge that the

> instrument's measurements are accurate according to a known standard, or are accepted as accurate by reputable scientists.[1]

1 (1980) 26 SASR 244, 253.

6.43 By implication, the court concluded that it would be extreme to establish the reliability of a software controlled device in a court of law by analysing the software code – the very software code that controlled the device and provided the evidence. The court considered that evidence from the operator of the device was sufficient for the trial court to assess the accuracy of the evidence. The appeal was dismissed.

6.44 Given these comments, it is understandable that the court reached the conclusions it did in *Mehesz v Redman (No 2)*. At issue was a self-contained device that was used by trained operators with suitable qualifications. On the basis that the readings from such devices were, at any time, not within the expected range, the suitably trained and qualified operators were expected to use their professional judgment to verify the reliability of the device before submitting the evidence for legal proceedings. In such a case, the court would not require the software code to be challenged.

6.45 The case of *Bevan v The State of Western Australia*[1] illustrates the approach taken when considering the admission of evidence from computers and computer-like devices. One of the grounds of appeal in this case was the admissibility of mobile telephone data in the form of text messages downloaded by a computer software program. An investigating police officer carried out two separate downloading operations using two separate tools, Cellebrite and XRY. At the beginning of the trial, counsel for the accused objected to the text messages being received into evidence. The trial judge held that the text messages were admissible. Questions were raised as to the reliability of the software and of the officer's correct use of it. The Court of Appeal concluded that the trial judge erred in law in admitting the text messages into evidence. This was because the officer did not explain the process of how he downloaded it in any detail at trial: it was the first time he had used the relevant software, and he did not have any formal training in its use. When considering the rebuttable presumption at common law as to the accuracy of 'notorious' scientific or technical instruments, Blaxell J said that 'when evidence from a new type of scientific instrument or process is adduced for the first time, there must be proof of its reliability and accuracy'.[2] He went on to say that:

> When specific evidence of the accuracy of a new instrument is required, this need not come from the manufacturer. It is sufficient that the expert who uses it can say that it is an instrument which is accepted and used by competent persons as a reliable aid in the carrying out of the scientific procedure in question, and that he so regards it.[3]

1 [2010] WASCA 101.
2 [2010] WASCA 101, [30].
3 [2010] WASCA 101, [31].

6.46 Blaxell J approved of the observations by White J[1] in *Mehesz v Redman (No. 2)* as noted above. He continued:

> To the above principles I add the obvious comment that a court will not be satisfied that an instrument was 'handled properly' on a particular occasion, if it does not understand what was required of the operator for this to be so. Detailed

> evidence as to the workings of the instrument need not be given … However, it is necessary that there be sufficient evidence for the court to apprehend what it was that the operator had to do in order to ensure an accurate result.[2]

1 *Mehesz v Redman (no 2)* (1980) 26 SASR 244 at [251]–[252].
2 [2010] WASCA 101, [33].

6.47 In essence, Blaxell J is saying that if the user of a smartphone can give evidence to demonstrate that he can use the smartphone, it follows that he is sufficiently knowledgeable to give evidence indirectly that the software code that controls the device is 'working properly', 'reliable' or 'accurate'. It is as if the software programs that form the device are irrelevant. Additionally, no attempt was made to define how software code can be determined to be 'working properly', 'reliable' or 'accurate'.

6.48 In *Bevan v The State of Western Australia*, the Court of Appeal heard a second appeal in the same case after a re-trial. The same argument arose regarding the method of downloading the data from the mobile telephone. There was a trial within a trial concerning the evidence of Detective Tomlinson. (Buss J referred to him as a First Class Constable, and set out his qualifications.[2]) Counsel for the appellant conceded that the witness was qualified to operate the equipment used to perform the download, but argued that he was not qualified to give evidence about the accuracy of the download material and the reliability of the material itself. In cross-examination, Detective Tomlinson explained he did not hold a certificate in relation to the Cellebrite and XRY software packages, but that he had been shown how to use them on about ten occasions. The following exchange took place regarding how the software worked:

> Q. Can you tell me how the Cellebrite package actually works.
>
> A. I don't understand the question.
>
> Q. How does it work? Explain to me, a layman, who knows nothing about Cellebrite, how it works.
>
> A. It extracts data from a telephone.
>
> Q. How? How does it do that?
>
> A. It uses software.
>
> Q. And how does that software work?
>
> A. I couldn't tell you.
>
> Q. What about the XRY?
>
> A. The same.
>
> Q. If you don't know how it works, how can you say its [sic] reliable?
>
> A. You'd have to ask the manufacturer.
>
> Q. Okay. I'm asking you. How can you say its [sic] reliable.
>
> A. I can't.
>
> Q. You can't. And, in fact, on one occasion that you used it in relation to the Nokia, it was unsuccessful.
>
> A. Yes, that's right.[3]

1 [2012] WASCA 153.
2 [2012] WASCA 153, [18]–[21], [105].
3 [2012] WASCA 153, [20], the last question and answer is at [106(g)].

6.49 In deciding to allow the evidence before the members of the jury, the trial judge said:

> The workings of the instrument need not be given and it seems to me that in this case the notes of the experienced officer, the evidence that this software is regularly used by him establishes the level of accuracy and in his notes at the time that he was – successfully used the program seems to me to meet the tests ... He was a trained, experienced and competent operator and the software was operated properly and, in those circumstances, in this case I think this evidence is admissible and I will allow it to be given by the qualified expert.[1]

1 [2012] WASCA 153, [201].

6.50 Pullin and Mazza JJA agreed the trial judge did not err in overruling the objection to the tendering of the text messages. In essence, because Detective Tomlinson was qualified as an expert, he could testify about the performance of the machines and the software. It was inferred that as an expert, he considered the process to be accurate, and that because he had performed such actions previously, the actions undertaken on this particular occasion were properly performed – even though the user of the program will not know that it is giving inaccurate results. There was no requirement for the Detective to understand how the software worked, or whether there were any problems with the software he used.[2] Pullin JA said: 'His evidence provided sufficient assurance that the results produced by the machines were reliable and accurate, because he (a trained operator of the machines) observed them to be so.'[3] But it does not follow that any operator of an electronic device will be able to detect if the device was malfunctioning in any way. As noted by Eric Van Buskirk and Vincent T. Liu:[4]

> There is a general tendency among courts to presume that forensic software reliably yields accurate digital evidence. As a judicial construct, this presumption is unjustified in that it is not tailored to separate accurate results from inaccurate ones.

1 As in the case of the death of Casey Marie Anthony in 2011, for which see Craig Wilson, 'Digital Evidence Discrepancies – Casey Anthony Trial' (11 July 2011) <www.digital-detective.net/digital-evidence-discrepancies-casey-anthony-trial/>; Tony Pipitone, 'Cops, prosecutors botched Casey Anthony evidence' (Clickorlando.com, 28 November 2012) <www.clickorlando.com/news/cops-prosecutors-botched-casey-anthony-evidence>; Jose Baez and Peter Golenbock, *Presumed Guilty: Casey Anthony: The Inside Story* (updated edn, BenBella Books 2013) 46, 180–183, 211, 346–348, 365, 368–371, 400, 426–428; Jess Ashton and Lisa Pulitzer, *Imperfect Justice: Prosecuting Casey Anthony* (William Morrow 2011) 105, 239, 277, 291–2, 298, 315.

2 [2012] WASCA 153, the rationale was set out at [66] and [67].

3 [2012] WASCA 153, [67].

4 Eric Van Buskirk and Vincent T Liu, 'Digital evidence: challenging the presumption of reliability' (2006) 1 Journal of Digital Forensic Practice 19.

6.51 They suggest there are two approaches to resolve the problem:

> One is through the proper application of scientific jurisprudence to questions of digital evidence and the other is through some combination of certain broad market and social corrections.

6.52 The important question is: *If the device was malfunctioning, how would the operator know?* More significantly, the question should be: *How would the malfunction manifest itself, if at all, and in a form evident to the operator?*

6.53 In the minority, Buss J considered that none of the relevant basic facts and circumstances were proved. The judge considered the applicable legal principles in detail.[1] He cited the relevant case law, and also extracts from *The Science of Judicial Proof* (3rd edn, 1937) by Professor Wigmore:

> Professor Wigmore enunciated three fundamental propositions applicable to evidence based on the use of a mechanical or scientific instrument constructed on knowledge of scientific laws:
>
> > 1. *The type of apparatus purporting to be constructed on scientific principles must be accepted as dependable for the proposed purpose by the profession concerned in that branch of science or its related art*. This can be evidenced by qualified expert testimony; or, if notorious, it will be judicially noticed by the judge without evidence.
> >
> > 2. *The particular apparatus used by the witness must be one constructed according to an accepted type and must be in good condition for accurate work*. This may be evidenced by a qualified expert.
> >
> > 3. *The witness using the apparatus as the source of his testimony must be one qualified for its use by training and experience* (§220).[2] (original emphasis)

1 [2012] WASCA 153, [111]–[129].
2 Para 111.

6.54 The judge continued:

> *Wigmore on Evidence* (Chadbourn Rev, Vol III, 1970) §795 states the requirements for the admissibility of evidence based on the use of scientific instruments, as follows:
>
> > What is needed, then, in order to justify testimony based on such instruments, is preliminary professional testimony: (1) to the *trustworthiness of the process* or instrument in general (when not otherwise settled by judicial notice); (2) to the correctness of the *particular instrument*; such testimony being usually available from one and the same qualified person. (original emphasis)[1]

1 [2012] WASCA 153, [112].

6.55 Buss J rejected the evidence of the Constable, partly because he was not qualified to comment of the software, and because the 'machines/software' were not so well-known that their accuracy may be assumed as a matter of common experience.[1] Evidence was required to demonstrate their accuracy. It followed that the State had to produce evidence from a suitably qualified expert of the trustworthiness of the machines and software in general, and of the correctness of the particular instruments for the purposes of downloading of data from mobile telephones.[2] Arguably, had the State produced sufficient evidence to convince a judge of the accuracy of the machines and software, it would not have been necessary to reply on the presumption. Notwithstanding this observation, the approach by Buss J is to be preferred. His brother judges appear to accept the astonishing conclusion that not having any knowledge of how a device works is irrelevant to the results of the analysis. In their approach, the work of software programmers is immaterial. Software code is not germane when determining causation. If this approach were accepted, no longer would decisions in legal proceedings be based on knowledge and systematic and scientific judicial inquiry.

1 This is a criterion that ignores how often people trust something that is untrustworthy simply because they are never tempted to challenge its results and scrutinize them with sufficient rigour to be able to tell whether they are correct.
2 [2012] WASCA 153, [132]–[139].

Mechanical instruments and computer devices

6.56 The discussion in this chapter focuses on the software code that provides instructions. In addition, the chapter concentrates on the software code in use by the user, as opposed to the operating system, which is also the subject of failure. In the case of firmware, which is software that is incorporated into hardware, the absence of visible programs does not mean that software is absent: the commentary in this chapter applies equally to this form of implementation of software.

The nature of software errors

6.57 It can be said that a computer can be both 'reliable' (but not infallible) and yet perform functions without the authority or knowledge of the owner or software writer. This may be when the code happens to execute in a way, because of a strange or unforeseen conjunction of inputs, which neither the owner nor the writer had imagined. For instance, one Jonathan Moore designed and produced forged railway tickets that were accepted by ticket machines controlled by computers. It took a ticket inspector to notice subtle differences in the colour and material of the ticket, which led to his arrest and prosecution for forgery.[1]

1 Tom Pugh, 'IT expert sentenced for rail ticket forgery', *The Independent* (London, 2 October 2009).

6.58 It is important to understand that programmers are aware of the limitations, as famously articulated by Ken Thompson:

> You can't trust code that you did not totally create yourself. (Especially code from companies that employ people like me). No amount of source-level verification or scrutiny will protect you from using untrusted code.[1]

1 Ken Thompson, 'Reflections on trusting trust', Turing Award Lecture (1984) 27 Communications of the ACM 761; further references Donald MacKenzie, *Mechanizing Proof Computing, Risk and Trust* (MIT Press 2004) 299, fn 1.

6.59 These comments are decidedly relevant, given that Thompson demonstrated how to create a C program fragment that would introduce Trojan Horse code into another compiled C program by compromising the C compiler. Thomas Wadlow explained this process as follows:

> For example, when compiling the program that accepts passwords for login, you could add code that would cause the [first] program to accept legitimate passwords or a special backdoor password known to the creator of the Trojan. This is a common strategy even today and is often detectable through source-code analysis.
>
> Thompson went one step further. Since the C compiler is written in the C programming language, he used a similar technique to apply a Trojan to the C compiler source itself. When the C compiler is compiled, the resulting binary

> program could be used to compile other programs just as before; but when the program that accepts passwords for login is compiled with the new compiler from clean, uncompromised source code, the backdoor-password Trojan code is inserted into the binary, even though the original source code used was completely clean. Source-code analysis would not reveal the Trojan because it was lower in the tool chain than the login program.[1]

1 Thomas Wadlow, 'Who must you trust?' (2014) 12 acmqueue Security 2.

6.60 Just because a person is in physical control of a computer or shop cash till, it does not follow that he will be aware whether it is working 'reliably', 'properly', 'consistently', 'correctly' or 'dependably'.[2] As indicated above, even the writer of the software will not be in such a luxurious position. It follows that the following comment by Kerr LCJ was not correct:

> In the modern world the presumption of equipment being properly constructed and operating correctly must be strong. It is a particularly strong presumption in the case of equipment within the control of the defendant who alone would know if there was evidence of incorrect operation or incorrect setting.[3]

1 Stephen Castell, 'Letter to the Editor' (1994) 10 Computer L & Secur Rep 158 pointed out that the observation by Lord Griffiths, at 387D, that a till was a 'computer … of the simplest kind' in *R v Shephard* [1993] AC 380 was, even at the time, an assumption that did not reflect the truth.

2 The use of the word 'dependability' is a global concept that subsumes attributes of reliability, availability, safety, integrity and maintainability, and 'reliability' provides for continuity of correct service: Algirdas Avižienis, Jean-Claude Laprie and others, 'Basic concepts and taxonomy of dependable and secure computing' (2004) 1 IEEE Transactions on Dependable & Secure Computing 11, 13.

3 *Public Prosecution Service v McGowan* [2008] NICA 13, [2009] N.I. 1, [20]; it is acknowledged that many standards in the safety critical community require some element of proof in the tools they use, such as evidence that the supplier tracks and corrects defects, for instance.

6.61 That software code is imperfect and remains so may be illustrated by the comments of an early pioneer in computing, the late Professor Sir Maurice V. Wilkes FRS FREng:[1]

> By June 1949 people had begun to realize that it was not so easy to get a program right as had at one time appeared. I well remember when this realization first came on me with full force. The EDSAC was on the top floor of the building and the tape-punching and editing equipment one floor below on a gallery that ran round the room in which the differential analyzer was installed. I was trying to get working my first non-trivial program, which was one for the numerical integration of Airy's differential equation. It was on one of my journeys between the EDSAC room and the punching equipment that 'hesitating at the angles of the stairs' the realization came over me with full force that a good part of the remainder of my life was going to be spent in finding errors in my own programs. Turing had evidently realized this too, for he spoke at the conference on 'checking a large routine'.

1 Maurice V Wilkes, *Memories of a Computer Pioneer* (MIT Press, 1985) 145.

6.62 This observation has been repeated many times since.[1] Professor Lloyd has expressed the view that the received wisdom is 'that all software contains defects'[2] – although he does not explain whether 'received wisdom' is based on evidence from technicians. Programmer errors are caused by a mix of novelty (applying software to previously unsolved problems), and the difficulty of the tasks software is required to perform, magnitude and complexity.[3] To address this problem, the approach of many of

the existing software safety standards is to define requirements for and put constraints on the software development and assurance processes.[4] Theodore A. Linden observed in 1976 that:

> It is more difficult to build a 50,000 line program than it is to write 1,000 programs that are each 50 lines long. This phenomenon leads to rapidly escalating costs for the development and maintenance of large software systems, and its leads to serious reliability problems due to the difficulty of adequately debugging and testing a large program.[5]

1 The reader might wish to begin with the following, which is only one of many articles by many eminent people: Les Hatton, 'Characterising the diagnosis of software failure' (2001) 18 IEEE Software 34.
2 Ian J Lloyd, *Information Technology Law* (7th edn, Oxford University Press 2014) 482.
3 B Littlewood and L Strigini, 'Software reliability and dependability: a roadmap' in A Finkelstein (ed.), *The Future of Software Engineering* (New York: ACM Press 2000) 177–88.
4 John McDermid and Tim Kelly, 'Software in safety critical systems: achievement and prediction' (2006) 2 Nuclear Future 34.
5 Theodore A Linden, 'Operating system structures to support security and reliable software' (1976) 8 ACM Computing Surveys (CSUR) 418.

6.63 Using the taxonomy of the provision of services, Algirdas Avižienis and colleagues have defined a 'correct service' as one where the service implements the system function. Its failure is an event that occurs when the service does not do what the function provides. This deviation is described as an 'error'. For instance, if the function when using an ATM is to dispense cash, and the ATM dispenses the correct amounts of cash, then there is a correct service, and the service is carried out in accordance with the function. If the amount of cash withdrawn from an ATM is greater or less than the amount keyed in, or no cash is provided, this is service failure that can be an error or fault. The authors go on to say:

> Since a service is a sequence of the system's external states, a service failure means that at least one (or more) external state of the system deviates from the correct service state. … In most cases, a fault first causes an error in the service state of a component that is a part of the internal state of the system and the external state is not immediately affected.
>
> For this reason, the definition of an **error** is the part of the total state of the system that may lead to its subsequent service failure. It is important to note that many errors do not reach the system's external state and cause a failure. A fault is **active** when it causes an error, otherwise it is **dormant**.[1]

1 Avižienis and others, 'Basic concepts and taxonomy of dependable and secure computing' 13; for additional discussions on this topic, see John Rushby, 'Critical system properties: survey and taxonomy' (1994) 43 Reliability Engineering and System Safety 189, and Donald MacKenzie, *Mechanizing Proof Computing, Risk and Trust* (MIT Press 2004) 337, fn. 16.

6.64 For instance, an ATM might provide a receipt that £100 has been withdrawn, but does not dispense the money. Given this set of facts, clearly a fault has occurred, because the sensors or the software code (or both) in the machine failed to detect the lack of movement of cash. The bank might provide a print-out of the machine's internal functioning that shows the balance of cash held in the machine before the transaction, and again after it. This proves very little. In the New York case of *Porter v Citibank, N.A.*,[1] a similar set of facts occurred. The customer used his card, but no money was dispensed. Employees of the bank testified that on average machines were

out of balance once or twice a week. From an evidence point of view, the information on the print-out is restricted to a single transaction. For the bank to prove that the machine actually dispensed £100 (and therefore the customer is lying), it is necessary for the bank to balance the ATM and report the results for the material time. The overall balance might indicate that it had gone down by £100. But the report might be inaccurate. This is because of a number of associated variables, such as (this is not an exhaustive list): the multiple layers of outsourcing, the fact that people cover up mistakes, and the fact that people rely on other people to be diligent in dual-control tasks. Equally, if the machine happens to overpay someone else by £100, the error will cancel out the previous error and the end result could not have been detected by human intervention either. Human cross checks may suggest that everything appears correct, but the system is failing repeatedly. A further reason for the machine to be in error is that a third party may have successfully inserted code to bypass the software in the machine, leaving the thief to recover the cash after the customer left the scene.[2]

1 123 Misc.2d 28, 472 N.Y.S.2d 582 (N.Y.City Civ.Ct. 1984).

2 Stephen Mason, 'Debit cards, ATMs and negligence of the bank and customer' (2012) 27 Butterworths Journal of International Banking and Financial Law 163; Maryke Silalahi Nuth, 'Unauthorized use of bank cards with or without the PIN: a lost case for the customer?' (2012) 9 Digital Evidence and Electronic Signature Law Review 95; Stephen Mason, 'Electronic banking and how courts approach the evidence' (2013) 29 Computer Law and Security Review 144.

6.65 For all these reasons, it is very hard to show that a computer is working 'properly', even for highly skilled professionals.[1] Part of the problem is that computers fail in discontinuous ways, which is a characteristic of discrete complexity, unlike most mechanical devices.

1 There is a technique called code verification, where code functionalities are verified as mathematical properties. But this process is time-consuming and limited. I owe this observation to Professor Seng.

Why software appears to fail

6.66 People across the world increasingly depend on computers and computer-like devices for mundane uses such as recording devices (cameras and recorders on mobile telephones), to critical uses such as lifesaving devices that control delicate medical equipment in hospitals to important infrastructural uses such as systems for the supply of gas, electricity and fuel, underground trains,[1] buses,[2] and financial software that assess risk in financial products.

1 The railway trains on London Underground's Jubilee line were being replaced from 2011. Many of the new trains failed and left passengers stranded for hours because of software failures: Dick Murray, 'Computer crash caused Jubilee line "meltdown"', *Evening Standard* (London, 9 November 2011) 11. This problem was also included in one of the series of six programmes by the BBC entitled 'The Tube' and broadcast during the spring of 2012.

2 A software problem meant the new London bus had to be run with its distinctive rear platform shut: 'New Routemaster bus starts running on London roads', *BBC News* (27 February 2012) <www.bbc.co.uk/news/uk-england-london-17173625>.

6.67 In the light of the ubiquitous nature of software, it is important to be aware that software code can function as intended by the programmer, but it can be the cause of failure. Alternatively, software code may fail to function in the way the designers intended, or it might continue to function but undertake actions that the designer did

not originally intend or instruct the device to undertake. Problems can occur for a number of reasons, such as where software code has a mistake, or because of improper installation.[1] A range of consequences might follow, such as failing air traffic control systems[2] and baggage handling systems in airports,[3] preventing couples from obtaining mortgages because of incorrect records,[4] dispensing more cash than is recorded via faulty software in ATMs,[5] miscalculating assets in family cases via software,[6] and causing injuries and deaths of people.[7]

1 One reason is because those people hired to undertake the work are not sufficiently qualified, as in *Robotic Vision Systems, Inc. v Cybo Systems, Inc.*, 17 F.Supp.2d 151 (E.D.N.Y. 1998).

2 Leonard Lee, *The Day The Phones Stopped The Computer Crisis-The What and Why of It, and How We Can Beat It* (Donald I. Fine 1991) ch. 7; Independent Enquiry, NATS System Failure 12 December 2014 – Final Report (13 May 2015), paras ES7 – ES10 <www.caa.co.uk/WorkArea/DownloadAsset.aspx?id=4294974241>.

3 Michael Schloh, *Analysis of the Denver International Airport baggage system* (Computer Science Department, School of Engineering, California Polytechnic State University 1996), available at <www5.in.tum.de/~huckle/schloh_DIA.pdf>; The Department of Homeland Security, Office of the Inspector General, *Lessons Learned from the August 11, 2007, Network Outage at Los Angeles International Airport (Redacted)* (OIG-08-58, May 2008); House of Commons Transport Committee, *The opening of Heathrow Terminal 5: Twelfth Report of Session 2007–08: Report, together with formal minutes, oral and written evidence* (HC 543, 3 November 2008).

4 Nicole Blackmore, 'Npower's error cost us our mortgage', *The Daily Telegraph* ('Your Money' London, 10 May 2014) 1, 3.

5 Tim Stewart, 'Huge queues as Tesco cash machine gives customers "free money"', *Evening Standard* (London, 18 August 2009), <www.standard.co.uk/news/huge-queues-as-tesco-cash-machine-gives-customers-free-money-6702682.html>; for other examples, see Stephen Mason, *When Bank Systems Fail: Debit cards, credit cards, ATMs, mobile and online banking: your rights and what to do when things go wrong* (2nd edn, PP Publishing 2014).

6 Owen Bowcott, 'Revealed: divorce software error hits thousands of settlements', *The Guardian* (London, 17 December 2015).

7 Donald MacKenzie, 'Computer-related accidental death: an empirical exploration' (1994) 21 Science and Public Policy 233.

Classification of software errors

6.68 The word 'bug' is a common term that is used in the information technology industry to describe a variety of issues.[1] When a technician uses this term, it can have a number of meanings.[2] Professor Thomas offered his view at a lecture he gave in 2015:[3]

> Different researchers and authors may describe faults as 'flaws', 'errors', 'defects', 'anomalies' or 'bugs' but they will almost always mean *functional* faults, which cause the software to crash or to give the wrong results.

1 It must be emphasized that there are a number of definitions of technical terms, but they are not dealt with in any detail in this text. For an insight as to how 'bugs' are dealt with in a contract between commercial entities, see *GB Gas Holdings Limited v Accenture (UK) Limited* [2010] EWCA Civ 912 and *Kingsway Hall Hotel Ltd v Red Sky IT (Hounslow) Ltd* [2010] EWHC 965 (TCC); in the software world, a 'bug' is also known as an undocumented feature, for which see David Lubar, *It's Not a Bug, It's a Feature!* (Addison-Wesley 1995).

2 The members of the team responsible for writing the following report did not use the term 'bug' when they meant 'error': Willis H Ware (ed.), *Security Controls for Computer Systems: Report of Defense*

Science Board Task Force on Computer Security – RAND Report R-609-1 (Published for the Office of the Secretary of Defense, R-609-1, Reissued October 1979).
3 'Should we trust computers?', a lecture given at Gresham College on 20 October 2015, available at <www.gresham.ac.uk/lectures-and-events/should-we-trust-computers>.

6.69 Lay people, not without some justification, consider the term 'bug' to be a cloak that hides the correct meaning, namely that what is being described is an error, flaw, mistake, failure, or fault in a software program or system.[1] Drawing from the work of Professor Ladkin, it is possible to classify most software errors into the following non-exhaustive categories:[2] human errors in coding and software development; software design or specification errors; unintended or unanticipated software interactions and input data flaws.

1 Causes of failure can also be categorized into human error, environment (including power outages or A/C failure), network failure, software failure and hardware failure: Bianca Schroeder and Garth A Gibson, 'A large-scale study of failures in high-performance computing systems' (2010) 7 IEEE Transactions on Dependable and Secure Computing 338.
2 Peter B Ladkin, *On Classification of Factors in Failures and Accidents* (Report RVS-Occ-99-02), available at <www.rvs.uni-bielefeld.de/publications/Reports/classification.html>.

Human errors in the software code

6.70 Notwithstanding the best software development tools that catch and identify coding errors, human errors in writing software code account for a large number of software errors. This problem is going to be exacerbated, given the increasing size of written codes. An example of human error in software code is that of *Mariner I*, the spacecraft that was sent to Venus and launched on 22 July 1962. The software code indicated that the booster had failed, and the rocket was destroyed on command from the control centre. In fact, the rocket was behaving correctly, and the computer system on the ground was at fault, partly because of a defect in the software, and partly because of a hardware failure. The error in the software arose because the person who wrote the software failed to include an overbar in the guidance equations.[1]

1 Peter G Neumann, *Computer Related Risks* (Addison-Wesley, 1995) 26–7 ('Here *R* denotes the radius; the dot indicates the first derivative – that is, the velocity; the bar indicates smoothed rather than raw data; and *n* is the increment. When a hardware fault occurred, the computer processed the track data incorrectly, leading to the erroneous termination of the launch.'); see also the explanation by the National Aeronautics and Space Administration report NSSDC ID: MARIN1, available at <http://nssdc.gsfc.nasa.gov/nmc/spacecraftDisplay.do?id=MARIN1>; for more detail of computers and the space age and an analysis of accidents (including this example), see Paul E Ceruzzi, *Beyond the Limits: Flight Enters the Computer Age* (MIT Press 1989).

6.71 Two further examples are the *Clementine* mission and the *Ariane 5* failure. The *Clementine* mission was a joint project between the Strategic Defense Initiative Organization and NASA. After the spacecraft left lunar orbit, a malfunction in one of the on-board computers on 7 May 1994 caused a thruster to fire until it had used up all of its fuel, leaving the spacecraft spinning at about 80 rpm with no spin control. The spacecraft remained in geocentric orbit and continued testing the spacecraft components until the end of mission.[1] In the case of the *Ariane 5* rocket failure in 1996, the disintegration of the rocket 40 seconds after launch was due to a software failure – because, in the words of Professor Les Hatton, 'the programmers had arranged the code such that a 64 bit floating point number was shoe-horned into a 16-bit integer'.[2]

As pointed out by Professor Ladkin, 'Code was reused from the Ariane 4 guidance system. The Ariane 4 has different flight characteristics in the first 30 seconds of flight and exception conditions were generated on both inertial guidance system (IGS) channels of the Ariane 5'.

1 <http://ntrs.nasa.gov/search.jsp?R=19980041408>; *Lessons Learned from the Clementine Mission* (Space Studies Board, National Research Council, National Academy Press 1997), available at <http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19980041408.pdf>.
2 Les Hatton, 'Ariane 5: A smashing success', *Software Testing and Quality Engineering* 2 (1999), pp. 14–16; Ariane 501 Inquiry Board report (4 June 1996), available at <http://esamultimedia.esa.int/docs/esa-x-1819eng.pdf> and <www.ima.umn.edu/~arnold/disasters/ariane5rep.html>; Charles C Mann, 'Why software is so bad' (2002) 38(b) Technology Review; Derek Partridge, *The Seductive Computer: Why IT Systems Always Fail* (Springer 2011) 99, fn. 6.
3 Peter B Ladkin, *The Ariane 5 Accident: A Programming Problem?* (Article RVS-J-98-02), available at <www.rvs.uni-bielefeld.de/publications/Reports/ariane.html>.

Failure of specification

6.72 The problem might not be in the software code, but with the specification,[1] such as with the loss of the *Mars Climate Orbiter* spacecraft in 1999.[2] On this occasion, the failure was not to use metric units in the coding of a ground software file. The thruster performance data used in the software application code entitled SM_FORCES (small forces) was in imperial units instead of metric units.[3] Roy Longbottom, Head of the Large Scientific Systems Branch of the Central Computer Agency, observed that:[4]

> When the software is first written and assembled, as for hardware, it usually undergoes a series of design quality assurance tests to ensure that the specification is met on facilities, performance and on physical source requirements. It is again fairly easy to check out the broad facilities provided but impossible to forecast and test for all possible modes of operation, combinations and sequences. One difference with hardware is that, the writing of comprehensive tests[5] for the software is often regarded as an overhead, whereas for hardware, comprehensive tests are written as a natural process for identifying constructional defects on all new equipment and for overcoming long term reliability problems. So, when software is first delivered, it is almost certain that the design will not be quite correct or some coding errors will be present.

1 For an example of the failure of a properly structured agreement that included what the customer wanted from the software, see *South West Water Services ltd v International Computers Ltd* [1999] Masons CLR 400.
2 In *Co-Operative Group (Cws) Ltd. (Formerly Co-Operative Wholesale Society Ltd.) v International Computers Ltd.* [2003] EWHC 1 (TCC), the case failed for lack of a contract, but the judge observed, at [260], that '… the initial efforts of ICL to try to meet the requirements of CWS as to when software was required were frustrated by the failure of CWS to specify precisely what its requirements were …'.
3 'Mars Climate Orbiter Mishap Investigation Board Phase I Report' (10 November 1999), available at <ftp://ftp.hq.nasa.gov/pub/pao/reports/1999/MCO_report.pdf>.
4 Roy Longbottom, *Computer System Reliability* (Wiley 1980) 71. This book may have been published in 1980, but remains true in the 21st century. Note Chapter 6 regarding faults.
5 Because of the discontinuous nature of software, the notion of a 'comprehensive test for software' does not exist, even in the high-integrity market. Testing every possible sequence of every possible input is not feasible.

6.73 It is a pervasive characteristic of software code that design will not be quite correct or coding errors will be present. The attitude taken by the National Aeronautics and Space Administration (NASA) towards software code was to consider it of secondary

importance. Although this view had changed over time, and a rigorous methodology was since implemented to provide for the better control and development of software code, NASA never produced error-free software code.[1]

1 Nancy G Leveson, 'Software and the challenge of flight control' in Roger D Launius, John Krige and James I Craig (eds), *Space Shuttle Legacy: How We Did It and What We Learned* (American Institute of Aeronautics and Astronautics 2013).

Unintended software interactions

6.74 Software code might function correctly, as intended by the programmer, but the interactions between individual components of the software code can be the cause of failure, because the designers of the system fail to account for all the potential interactions. This is because the potential number of defects in software relates not only to the components (lines of code), but also to the number of ways in which they interact – the number of interactions increases faster than the number of components, thus making large systems with many components proportionally harder to get right. As the work of Bianca Schroeder and Garth A. Gibson demonstrates, the more complex the system becomes, the more likely it is that different types of failure will occur,[1] and the number of reasons that complexity causes failure also increases.[2] To put the problem into perspective, it is necessary to understand not the number of defects per device but the proportion of design decisions that contain defects, which might be termed a frequency.[3] A typical design decision in software looks like this:

```
if some-condition-I-have-decided-when-I-designed-the-software
then
                do something
otherwise
                do something else
```

1 Schroeder and Gibson, 'A large-scale study of failures in high-performance computing systems'.
2 For the same discussion in 1986, see Rudolph J Peritz, 'Computer data and reliability: a call for authentication of business records under the federal rules of evidence' (1986) 80 Northwestern University Law Review 965, 990–9; Stephen Mason and Timothy S Reiniger, '"Trust" between machines? Establishing identity between humans and software code, or whether you know it is a dog, and if so, which dog?' (2015) 21 CTLR 135–48; for a specific case study, see Sivanesan Tulasidas, Ruth Mackay, Pascal Craw, Chris Hudson, Voula Gkatzidou and Wamadeva Balachandran, 'Process of designing robust, dependable, safe and secure software for medical devices: point of care testing device as a case study' (2013) 6 Journal of Software Engineering and Applications 1.
3 Nobody is certain how many defects occur per lines of code or number of design decisions, but for a good discussion, see McDermid and Kelly, 'Software in safety critical systems'.

6.75 This means, illustrating the point with this simple example, that each design decision creates at least two choices for the software to handle, and within the 'do something' bits, further design choices will have to be made. This demonstrates that in software, a very few decisions rapidly creates a far more complex thing than humans can reliably analyse and be confident they have made the right decisions, in even a modest fraction of the possible cases.[1] Since there are typically thousands of design decisions in the software for even relatively small products, there will be hundreds of defects in the final products – Professor Pham suggests 'that as software projects become larger, the rate of software defects increases geometrically'.[2] An average defect

level of one to five defects per thousand lines of code could translate into hundreds if not thousands of defects for devices that have several hundred thousand to a million or more lines of code.[3] This is the typical size of most software that controls aircraft,[4] motor vehicles and many other common systems. The user is affected by how often the software fails, or how likely it is that in a particular occasion the software failed – a probability rather than a frequency, not by how many defects there are. This is because one defect may cause failures frequently, and another defect cause failures only very seldom.

1 I owe this analysis to Professor Harold Thimbleby.

2 Exponential is a more precise term than geometric: Hoang Pham, *System Software Reliability* (Springer 2000) 2. The software included in motor vehicles (called 'electronic control units' in the trade) is increasing in numbers, and has elaborate structures, all of which can lead to malfunctions that can cause death if the software is not properly tested: J. Mössinger, 'Software in automotive systems', (2010) 27 IEEE Software 92; Stephen Mason, 'Vehicle remote keyless entry systems and engine immobilisers: do not believe the insurer that they are perfect' (2012) 28 Computer Law and Security Review 195 in which it was predicted that the number of vehicles with keyless entry systems being stolen would increase, for which see Carnegie Menon, 'Hi-tech thieves add computers to crowbars', *The Guardian* (London, 25 June 2016) 49 and 'Is your car the most stolen model in England and Wales?' at <www.theguardian.com/money/2016/jun/25/hi-tech-thieves-keyless-car-crime-electronic-security>.

3 William Guttman, professor of economics and technology at Carnegie Mellon University, is of the view that the figure is nearer 30 errors per 1,000 lines of code on average: Alorie Gilbert, 'Newsmaker: Fixing the sorry state of software', *CNET News* (9 October 2002) (this item no longer seems to be available online).

4 On 2 June 1994, Chinook helicopter ZD 576 crashed on the Mull of Kintyre. The RAF Board of Inquiry held the pilots to be negligent. However, some considered that the installation of a Full Authority Digital Engine Control (FADEC) system was to blame, as described in detail in *RAF Justice* (Computer Weekly) <http://cdn.ttgtmedia.com/rms/computerweekly/DowntimePDF/pdf/rafjust.pdf>; 'Chinook crash: critical internal memo on software flaws', *Computer Weekly* (4 June 2009) <www.computerweekly.com/news/2240089594/Chinook-crash-critical-internal-memo-on-software-flaws>; the decision of the RAF Board of Inquiry was subsequently reversed: *The Mull of Kintyre Review* (HC Paper 1348, 2011) <www.gov.uk/government/uploads/system/uploads/attachment_data/file/247259/1348.pdf>.

6.76 This issue is further magnified by what are called 'legacy' systems. For instance, the computer systems used by airlines are very complex. There are a number of reasons: airlines introduced computer systems in the 1950s; as airlines merge, or take over other airlines, they might combine or adopt the computer systems they have inherited. Over time, as new functions are added, this process has created systems of great complexity. The banking sector has the same problem. Replacing such systems is not an easy decision, because it would take a considerable amount of money and time, and it is doubtful whether any IT firm has sufficient skills and knowledge to provide all the software needed for a complete replacement.[1]

1 'All systems stop: why big firms like Delta find it so hard to eliminate glitches from their IT systems', *The Economist* (London, 13 August 2016) (from the print edition) at <www.economist.com/news/business/21704842-why-big-firms-delta-find-it-so-hard-eliminate-glitches-their-it-systems-all-systems>.

6.77 One example of such a failure is the loss of the *Mars Polar Lander* and *Deep Space 2* missions. The loss of the spacecraft the failure is recounted in the NASA report:

> 7.7.2 Premature Descent Engine Shutdown

> FAILURE MODE DESCRIPTION
>
> A spurious signal, generated when the landing legs are deployed at an altitude of about 1500 meters, can cause premature descent engine shutdown when the lander is 40 meters above the surface.
>
> …
>
> The touchdown sensors characteristically generate a false momentary signal at leg deployment. This behavior was understood and the flight software was required to ignore these events; however, the requirement did not specifically describe these events, and consequently, the software designers did not properly account for them. The resulting software design recorded the spurious signals generated at leg deployment as valid touchdown events. When the sensor data were enabled at an altitude of 40 meters, the engines would immediately shut down. The lander would free fall to the surface, impacting at a velocity of 22 meters per second (50 miles per hour), and be destroyed.[1]

1 *Report on the Loss of the Mars Polar Lander and Deep Space 2 Missions* (JPL Special Review Board, 22 March 2000, JPL D-18709).

6.78 Professor Leveson describes this as a *component interaction accident*,[1] where an accident arises because of the interactions between the components of a system, rather than in the failure of any individual component. This is an example of incorrect software requirements, specifically of incorrect dependencies between components: the assumptions made in one element become an important part of the context of the requirements for some other part.[2] This illustrates the point that software itself is neither acceptably safe nor unacceptably unsafe; it is the operation of the software that might be called safe or unsafe.

1 Nancy G Leveson, *Engineering a Safer World: Systems Thinking Applied to Safety* (MIT Press 2011) 8, 49, 66–67.

2 For examples of other accidents, see Trevor Kletz, Paul Chung, Eamon Broomfield and Chaim Shen-Orr, *Computer Control and Human Error* (Gulf Professional Publishing 1995).

6.79 Consider a practical example. The display on the screen has a meaning, and if that meaning is not veridical, then an accident may result. Where the moon rising over the horizon causes a system to interpret it as a massive ICBM launch, semantic safety is violated: that is, the display (it might be a warning signal or something else) was not veridical. This problem has been linked to the possibility that a nuclear war has been averted by human intervention despite computer warnings of imminent attacks at least twice.[1]

1 I owe this suggestion to Professor Peter Bernard Ladkin. For the incident where software code made it appear the Soviet Union had launched an assault of nuclear missiles on the United State of America, see Donald MacKenzie, *Mechanizing Proof Computing, Risk and Trust* (MIT Press 2004) 23–4 and Eric Schlosser, *Command and Control* (Penguin 2014) 253–4; for an incident where software code made it appear there was a missile attack by the United States of America against the Soviet Union, see Ron Rosenbaum, *How the End Begins: The Road to a Nuclear World War* III (Simon & Schuster 2011) 7, 225–6, 248; Pavel Aksenov, 'Stanislav Petrov: The man who may have saved the world', *BBC News* (26 September 2013).

6.80 It should be observed that the increasing use of machine-learning systems complicates this issue, because the software code is instructed to make further decisions when running, which increases the complexity. In addition, the veridicality of machine-learning systems like neural nets cannot be easily understood or verified.[1]

1 I owe this point to Dr Michael Ellims and Professor Martyn Thomas, CBE, FREng.

Input data flaws

6.81 In addition, there are also what is known as 'input-data flaws', meaning that the data entered into the machine was not correct, thus ensuring the information coming out is also incorrect – colloquially known as 'garbage-in-garbage-out'. In a well-designed system, the software should check, insofar as that is possible, that the input data is wrong, corrupted or unexpected, and subject the output to a warning, perhaps via the user interface. This is a common problem in fairly simple systems such as databases, even in critical uses as in the medical field. Staff who enter data do not realize either the likelihood or the gravity of introducing erroneous data; and staff obtaining data from the database may not realize how likely the data are to be erroneous. These two factors combine into making the use of the software effectively cause wrong behaviour despite the software operating correctly.

Operational errors

6.82 Another manifestation of human error would be operational error. Professor Leveson observed that it is 'often very difficult to separate system design error from operator error: In highly automated systems, the operator is often at the mercy of the system design and operational procedures'.[1] The accuracy of this comment applies to virtually every automated system that includes computers and software code, and has, indirectly, caused significant loss of life. For instance, 'user interface errors' have been blamed for several aviation accidents, where the pilot as the user did not do anything wrong, but did not know the correct way to do what he wanted to do. Even in situations where people are part of a controlled and trained user community, such as ambulance controllers or air traffic controllers – human error rates in many tasks are high enough to stress systems in ways that are unpredictable. Examples of such situations in high stress industries such as the emergency services industry and the aviation industry are further explored in the rest of this chapter.

1 Leveson, *Engineering a Safer World* 39.

The development, maintenance and operation of software

6.83 As general purpose computing systems have become more powerful and flexible, users have devised new uses in ways that the systems developers never envisaged. This, coupled with the increase in complexity and the speed at which computers work, especially in modern automated systems, means that developers can never completely anticipate how users will use their products, or how their products will interact with other products and software. Even where the developers have tested their products in the ways that most users use them (and possibly fail to test them against less conventional methods of use), developers may subsequently provide upgrades that provide more functions, or issue updates to remedy any defects that have been found. In doing so, the developers will have modified the software and its operating conditions. Such changes will result in new modes of operation that have not been previously tested, causing the users to encounter defects they have not previously experienced. This problem is compounded when complex systems such as banking systems are linked to each other.

6.84 While it might appear that exhaustive testing could be the answer to this problem, it is impractical, does not necessarily work, and there is no workable theory that would constitute what constituted an adequate test. Professor Thomas notes that 'The main way that software developers assure the quality of their work is by running tests, even though computer scientists have been saying for the past forty years that testing can never show that software is secure or correct'.[1] For even relatively small systems, the number of possible test cases required for comprehensive testing is enormous. It is also not always certain whether or not the test has passed or failed, and it is necessary to repeat the tests after any software change. Furthermore, a single test case can only expose a system to a very specific set of conditions and data values. The number of variations is, in practical terms, unbounded because a robust test must consider, among others, different data values, the number of simultaneous jobs running, the system memory configuration, the hardware configuration, all of the connected devices or systems, the operators' actions, user errors, data errors, device malfunctions, and so forth. However, because testing is a complex affair does not mean that testing should not be carried out. This is so especially when people can be killed and injured, as in the case of the sudden unintended acceleration problems experienced by owners of some modern motor vehicles, which operate with electronic control systems. Michael Barr, in giving evidence as the expert witness for the plaintiffs in the trial of *Bookout v Toyota Motor Corporation Case*, gave the following in oral testimony:

> [Toyota] didn't [have] a formal safety process like the MIRSA, the big book. They don't follow a recipe for making a safe system.
>
> They also have the defect that they didn't do peer reviews on the operating system code or the monitor CPU codes. And here, ultimately, it comes down to resources. Toyota did not put people and time behind checking up on the suppliers who were supplying this critical software [for their vehicle electronic control systems]. The operating system at the heart of this main CPU and this and second CPU that's doing the monitoring.[2]

1 Martyn Thomas, 'Technology, security and politics' (2016) 25 SCSC Newsletter 53.

2 No. CJ-2008-7969 (Reported by Karen Twyford, RPR): examination and cross examination of Michael Barr 14 October 2013, 80, available at <www.safetyresearch.net/Library/Bookout_v_Toyota_Barr_REDACTED.pdf>.

Developmental issues and software errors

6.85 In examining the nature of a software fault, even at a time when software was less complex than now, Professor Randell and his colleagues made the following astute observation:

> A detected error is only a symptom of the fault that caused it, and does not necessarily identify the fault. Even where the relationship between the fault and the detected error appears obvious, it will be found that many other possible faults could have caused the same error to be detected.[1]

1 B Randell, P Lee and P C Treaven, 'Reliability issues in computing system design' (1978) 10 ACM Computing Surveys (CSUR) 126, 127.

6.86 Professor Randell also commented that 'What is significant about software faults is, of course, that they must be algorithmic faults stemming from unmastered complexity in the system design'.[1] This is a telling observation, in that the primary

source of software errors lies in its development process. There are numerous issues in the development of software that will generate errors, including but not limited to the speed that a developer is required to work to write proprietary software within the contractual time-frame, the consistent failure within the industry to provide for suitable quality control procedures, the creation of a climate of fear to suppress concerns relating to errors and safety,[2] and the insufficiency or lack of knowledge that programmers may have of the domain in which the software is to work (for instance, the programmer might be knowledgeable about mathematics, but have no knowledge of how acceleration systems work in motor vehicles[3]). Unrealistic estimates of how long it will take to write and test software also undermine accuracy, which means that those responsible for writing software code will not have the time or resources to be comprehensive in developing the software.[4] It is also necessary to have a comprehensive design that has been subjected to peer review that should precede any coding. Often, the writing of lines of code remains the ready, and easily quantifiable, measure of progress, which means that writing code starts much too quickly, and too little emphasis is placed on good design.

1 Randell, Lee and Treaven, 'Reliability issues in computing system design' 127.

2 Nancy G Leveson, 'Technical and managerial factors in the NASA Challenger and Columbia losses: looking forward to the future', in Daniel Lee Kleinman, Karen A Cloud-Hansen, Christina Matta and Jo Handelsman (eds), *Controversies in Science and Technology Volume 2: From Climate to Chromosomes* (Mary Ann Liebert Press 2008); for a legal response to this problem, see Richard Warner and Robert H Sloan, 'Vulnerable software: product-risk norms and the problem of unauthorized access' (2012) 45 University of Illinois Journal of Law, Technology & Policy.

3 Michael Ellims, 'On wheels, nuts and software', 9th Australian Workshop on Safety Related Programmable Systems (SCS'04) in Brisbane, 2.1, available at <http://crpit.com/abstracts/CRPITV47Ellims.html>.

4 This is not a recent phenomenon – even in 1976 it could be said that 'debugging and testing often account for half the cost of a program': Linden, 'Operating system structures to support security and reliable software', 410–11; and more recently, Robert N Charette, 'Why software fails' (2005) 42 IEEE Spectrum, available at <http://spectrum.ieee.org/computing/software/why-software-fails>; Partridge, *The Seductive Computer*; W Wayt Gibbs, 'Software's chronic crisis', *Scientific American* (September 1994) 86.

6.87 This is not to say that all software programmers are incompetent or that they do not wish to undertake work of a high quality. In their writings about software errors, Algirdas Avižienis and colleagues define 'human-made faults'[1] as including faults of omission, and wrong actions that lead to faults of commission. 'Human-made faults' are, in turn, divided into malicious faults and 'nonmalicious' or guileless faults. These faults can be introduced during the development of the system by a developer or during use by an external third party. Guileless faults can be classified as faults due to mistakes, and deliberate faults that are brought about because of bad decisions – usually caused when choices are made to accept having less of one thing in order to get more of something else, for instance, to preserve acceptable performance, or because of economic considerations. Developers who commit such faults may deliberately violate an operating procedure without understanding the consequences of their action, as illustrated by the organizational causes that led to the loss of the space shuttle *Columbia* and its seven-member crew.[2]

1 Avižienis and others, 'Basic concepts and taxonomy of dependable and secure computing' 15–18.

2 Columbia Accident Investigation Board, *Report Volume 1* (August 2003) 9.

6.88 The main part of the problem is that writing software used to be an exceedingly difficult and challenging field, and the methods used by management to control quality are not necessarily the most competent that can be used. However, writing software is now much easier. Advanced development environments generate code automatically, although writing software to perform complex functions that works well in all circumstances remains exceedingly difficult and challenging. Many amateurs have had the experience of being able to build software that achieves impressive effects with very little effort. This may well lead them to believe that because they find it easy to program a simple videogame or puzzle-solver (whose failures do not matter and will probably go unnoticed), or some simple program that seems reliable enough for their personal, everyday use, then building complex software systems that are correct must be just as easy. A further barrier arises when the organization is collectively incompetent.[1] This in turn means that inherent problems in software used in large organizations may not be identified for a long time. For instance, in 2003, Oates Healthcare began to use a new software product that was written for the company. At the time it began to be used, it was not known that the code written by the programmer was defective, in that it failed to calculate overtime for employees correctly. The problem was identified when a previous employee took legal action against the company five years after the software was used. As a result of discovering this problem, the company had to undertake two exercises. First, the simple solution was to write new software code to permit the software program to begin calculating overtime correctly from the point in time that the software was amended. Second, because the changes to the software were not capable of affecting the previous calculations, the previous records had to be re-calculated manually. Apparently there were over 10 million records that needed to be recalculated. A software project can fail partly because of a combination of the failure of management, an unrealistic time frame to develop the software, and a failure to develop and test software properly. There are many examples of such failure, and more importantly, some failures do not come to light until after the project is complete.[3]

1 As in the example of the failure of the AAS system: Office of Inspector General, 'Audit report advance automation system federal aviation administration', Report Number: AV-1998-113 (15 April 1998), available at <https://www.oig.dot.gov/sites/default/files/av1998113.pdf>.

2 Phil Simon, *Why New Systems Fail: Theory and Practice Collide* (AuthorHouse 2009) 7–9.

3 Robert L Glass, *Software Runaways: Lessons Learned from Massive Software Project Failures* (Prentice Hall PRT 1998) xiii–viv; Leonard Lee, *The Day The Phones Stopped: The Computer Crisis-The What and Why of It, and How We Can Beat It* (Donald I. Fine 1991); Nancy G Leveson, 'Role of software in spacecraft accidents' (2004) 41 Journal of Spacecraft and Rockets 564, also available at <http://sunnyday.mit.edu/papers/jsr.pdf>.

Increasing the risk of errors through modification of software

6.89 Software typically goes through modification cycles, called updates or upgrades to fix existing errors in code or enhance or improve software functionality. One of the major causes of software failure is that as software code is modified, each modification is capable of increasing the risk of failure. Some of the changes that are meant only to fix errors may create another one, resulting in a greater or smaller probability of failure. Where a vendor releases a significant number of new features or a major redesign, there is, typically, a sudden increase of the probability of failure, after which, the risk is reduced once further error updates begin to resolve the errors discovered, thus reducing the risk again over time.

6.90 It is useful to observe that when safety-related software code is modified, there is usually documentation to explain how the risk has been reduced, although this is only in the case of dangerous failures, and not necessarily all failures. By way of example, consider the case of *Saphena Computing Limited v Allied Collection Agencies Limited* in which Mr Recorder Havery QC commented:

> In the present case, on the other hand, once the software is fit for its purpose, it stays fit for its purpose. If by any chance a flaw is discovered showing that it is unfit for purpose (which is hardly likely after prolonged use)[1] there is a remedy in damages against the supplier, if solvent, until the expiry of the period of limitation.[2]

1 Professor Thomas has indicated that even in 1995 there was plenty of evidence that this was not correct.
2 [1995] FSR 616, 639.

6.91 The problem with this remark is that proprietary software code can be (and indeed often is) affected by updates, which means it does not necessarily stay 'fit for purpose'. Flaws can become manifest at any time, and some flaws can remain for years, which means if they are detected by a malicious person or state agency, they can be manipulated for purposes other than what users intend. There is a more fundamental flaw in this statement. If the software is used unchanged for a different purpose, which may be no more than the original purpose but applied to different data, it may still fail.

6.92 This is illustrated in the Heartbleed exposé.[1] Cryptographic protocols are used to provide for the security and privacy of communications over the Internet, such as the World Wide Web, email, instant messaging and some virtual private networks. The current protocol is called the Transport Layer Security (TLS). To implement this protocol, a developer will use a cryptographic library. One such library, which is open sourced, is OpenSSL. In 2011, a doctoral student wrote the Heartbeat Extension for OpenSSL, and requested that his implementation be included in the protocol. One of the developers (there were four) reviewed the proposal, but failed to notice that the code was flawed. The code was included in the repository on 31 December 2011 under OpenSSL version 1.0.1. The defect allowed anyone on the Internet to read the memory of any system that used the flawed versions of the OpenSSL software. It was possible for a hacker using this flaw to steal user names and passwords, instant messages, emails and business documents. No trace would be left of the attack. The attack did not rely on access to privileged information or credentials such as username and passwords. Taking into account the length of exposure, the ease by which it can be exploited, the fact that an attack does not leave a trace, and that it is estimated to have affected up to two-thirds of the Internet's web servers, this weakness was taken seriously. On 7 April 2014, the day the Heartbleed vulnerability was publicly disclosed, a new version that applied a fix to the flaw was released on the same day.

1 Jane Wakefield, 'Heartbleed bug: what you need to know', *BBC News Technology* (10 April 2014) <www.bbc.co.uk/news/technology-26969629>; Brian Krebs, *Heartbleed Bug: What Can You Do?* (14 April 2014) <http://krebsonsecurity.com/2014/04/heartbleed-bug-what-can-you-do/>; <https://en.wikipedia.org/wiki/Heartbleed>. A more important error was discovered in GNU Bash in September 2014, for which see 'Bourne-Again Shell (Bash) Remote Code Execution Vulnerability' (Original release date: 24 September 2014; last revised: 30 September 2014), at <https://www.us-cert.gov/ncas/current-activity/2014/09/24/Bourne-Again-Shell-Bash-Remote-Code-Execution-Vulnerability>.

6.93 Software can also be affected by changes in the environment, such as the operating system or other components, rather than any specific application, although it is necessary to distinguish between modification of software in situ and the reuse of software in an environment that is presumed to be similar. An example is the *Ariane 5* incident, where the malfunction arose from a changed environment and assumptions that were poorly understood, rather than a defect in the original development. Where the software is modified in situ, the environment does not change; where software is re-used in an environment that is presumed to be similar, the software has not changed, but the environment has. The results in either case are that there may be a mismatch where there was none before.

6.94 Generally speaking, programmers who modify someone else's code often do not fully understand the software, and may also be less well trained than the people who wrote it. Software can be relied upon to produce verifiably correct results, but to have such a degree of certainty, it is necessary to be assured that the operating conditions remain identical and that nothing else malfunctions. Peter G. Neumann has indicated that even though the utmost care and attention might be devoted to the design of a system, it may still have significant flaws.[1] This was illustrated in a 1970 report edited by Willis H. Ware.[2] Now freely available, the authors noted, under 'Failure Prediction' within section V System Characteristics, that:

> In the present state of computer technology, it is impossible to completely anticipate, much less specify, all hardware failure modes, all software design errors or omissions, and, most seriously, all failure modes in which hardware malfunctions lead to software malfunctions. Existing commercial machines have only a minimum of redundancy and error-checking circuits, and thus for most military applications there may be unsatisfactory hardware facilities to assist in the control of hardware/software malfunctions. Furthermore, in the present state of knowledge, it is very difficult to predict the probability of failure of complex hardware and software configurations; thus, redundancy [is] an important design concept.

1 Neumann, *Computer Related Risks*, 4; see his text generally for this topic.
2 *Security Controls for Computer Systems: Report of Defense Science Board Task Force on Computer Security – RAND Report R-609-1* <www.rand.org/pubs/reports/R609-1/index2.html>.

6.95 The authors of the report went on to observe the following in Part C, Technical Recommendations:

> (a) It is virtually impossible to verify that a large software system is completely free of errors and anomalies.
>
> (b) The state of system design of large software systems is such that frequent changes to the system can be expected.
>
> (c) Certification of a system is not a fully developed technique nor are its details thoroughly worked out.
>
> (d) System failure modes are not thoroughly understood, cataloged, or protected against.
>
> (e) Large hardware complexes cannot be absolutely guaranteed error-free.

Security vulnerabilities

6.96 Software vulnerabilities are software errors generally hidden from view. While they generally cause users no harm, they may be exploited by state security services, malicious hackers and professional thieves for various advantages, including theft of personal data (for sale), control of vulnerable systems, drug smuggling,[1] blackmail and other forms of financial gain. The market in selling packets of software code known as 'exploits' has become significant. Legitimate business may sell a vulnerability in a software code to business and government agencies, and hackers may sell a vulnerability to anyone who will buy them.[2] These vulnerabilities, particularly those against whom there are no pre-existing defences – known as 'zero day exploits' – may be exploited, whether legally or illegally, for criminal investigation as well as for cyber espionage purposes, including the violation of confidentiality (stealing information); availability (denial of service for political intimidation or blackmail) and of integrity (corrupting information to steal from banks or to cause an embedded computer system to cause accidents).

1 Hackers deployed to facilitate drugs smuggling, Intelligence Notification 004-2013, June 2013, Europol Public Information, available at <www.europol.europa.eu/category/publication-category/cyber-bits>.

2 'Cyber-security: the digital arms race', *The Economist* (London, 30 March 2013), at <www.economist.com/news/business/21574478-market-software-helps-hackers-penetrate-computer-systems-digital-arms-trade>.

6.97 To address these vulnerabilities, software vendors often, but not always, issue 'Security Patches' regularly each month (sometimes referred to as 'Software Updates' to conceal the nature of the up-date) in recognition of the failure of their software. By way of example, Microsoft issued 679 security bulletins between 4 January 2000 and 8 December 2009 (of which almost 300 were deemed 'critical'), 34 of which were released in October 2009 alone. With certain exceptions, software vendors issue regular updates, sometimes hiding or incorporating the nature of an up-date under the guise of a new version of the software.[1] Yet these may give rise to more problems. For instance, an important security weakness was discovered in relation to the distribution of software patches (which, ironically, was put in place to address security weaknesses). This meant that attackers who receive the patch first might compromise vulnerable hosts who have yet to receive the patch.[2]

1 A list of bulletins is available at <https://technet.microsoft.com/en-us/security/bulletins>.

2 David Brumley, Pongsin Poosankam, Dawn Song and Jiang Zheng, 'Automatic patch-based exploit generation is possible: techniques and implications' (Carnegie Mellon University, Carnegie Institute of Technology, Mellon Department of Electrical and Computing Engineering, Paper 44, 2008), available at <http://repository.cmu.edu/ece/44>.

6.98 Software security vulnerabilities are particularly pertinent to businesses and industries that operate or rely on digital-security infrastructures. For these industries, there are other issues to consider. The first is whether the design of the security protocol is robust. An example of a failure in this category is with banking systems,[1] and although a design can be modified, at best it is only possible to take a provisional view in respect to this point, because designs constantly change, and are therefore liable to failure. The second is whether the security protocol is implemented properly. For instance, a number of ATMs were tested around Cambridge in the UK, and it was

found that the nonce generation was predictable. A nonce is supposed to be a unique object in a protocol, a one-time 'security code', but it was found out that some ATMs were using a small supply of tokens as nonces and reusing them in a predictable order, thereby compromising their security.[2]

1 Steven J Murdoch, Saar Drimer, Ross Anderson and Mike Bond, 'Chip and PIN is broken', *31st IEEE Symposium on Security and Privacy* (IEEE Computer Society 2010) 433–46, available at <www.cl.cam.ac.uk/research/security/banking/nopin/oakland10chipbroken.pdf>; Steven J Murdoch, 'Reliability of Chip & PIN evidence in banking disputes' (2009) 6 Digital Evidence and Electronic Signature Law Review 98.

2 M. Geuss, 'How a criminal ring defeated the secure chip-and-PIN credit cards', *arstechnica* (20 October 2015), available at <http://arstechnica.com/tech-policy/2015/10/how-a-criminal-ring-defeated-the-secure-chip-and-pin-credit-cards/>; Mike Bond, Omar Choudary, Steven J Murdoch, Sergei Skorobogatov and Ross Anderson, 'Chip and Skim: cloning EMV cards with the pre-play attack', a paper presented to Cryptographic Hardware and Embedded System (CHES) 2012, in Leuven, Belgium, September 2012, available at <http://sec.cs.ucl.ac.uk/users/smurdoch/papers/oakland14chipandskim.pdf>; Houda Ferradi, Rémi Géraud, David Naccache and Assia Tria, *When Organized Crime Applies Academic Results: A Forensic Analysis of an In-Card Listening Device*, available at <http://eprint.iacr.org/2015/963.pdf>.

6.99 Furthermore, security may be associated with safety. If there is a safety-related system with security vulnerabilities, it is possible for the safety functions in the system to be deliberately subverted and give rise to a safety issue. For instance, the nuclear industry has developed a draft international standard for safety and security.[1] The vital problem in this area, which nobody has solved, is that while updates of safety functions in code that control nuclear reactors are slow, deliberate, and highly analytical, updates for security purposes have to be rapid, to forestall anticipated attempts via zero-day exploits. These two modi are obviously incompatible.

1 Caroline Baylon, with Roger Brunt and David Livingstone, *Cyber Security at Civil Nuclear Facilities: Understanding the Risks*, Chatham House Report (The Royal Institute of International Affairs, September 2015), available at <www.chathamhouse.org/publication/cyber-security-civil-nuclear-facilities-understanding-risks>.

6.100 It follows that software security vulnerabilities expose them to manipulations without the authority or knowledge of the software vendor.[1] Many of the vulnerabilities arise specifically from the errors in the original implementation. For instance, it might be possible for a person to control another owner's computer as part of a botnet[2] or enter the control system of an aircraft in flight via the in-flight entertainment system.[3]

1 The Trojan horse problem was recognized very early, for which see Linden, 'Operating system structures to support security and reliable software' 422–4.

2 Sanjay Goel, Adnan Baykal and Damira Pon, 'Botnets: the anatomy of a case' (2005) 1 Journal of Information System Security 45.

3 See the *Applicant for a Search Warrant in the case of Chris Roberts at the United States District Court for the Northern District Court of New York* Case number 5:15-MJ-00154 (ATB) dated 17 |April 2015, [18]–[19], available at <www.wired.com/wp-content/uploads/2015/05/Chris-Roberts-Application-for-Search-Warrant.pdf>.

6.101 At this point, the reader might consider that such problems can be solved fairly easily – by the introduction of anti-virus software (this is not to imply that all attacks are by the use of malicious software). But it must be understood that the fundamental nature of most anti-virus software limits its effectiveness – and the anti-virus software itself might not be error-free. A sophisticated attacker will have access to all the types

of anti-virus software, and he will program round the detection mechanisms and test his code against the anti-virus systems to ensure it is not detected.[1] Most anti-virus software is reactive, in that it searches for known threats. As such, anti-virus software is far from perfect. It fails to stop some malicious software[2] and should not be relied upon as the sole method of securing a computer. Indeed, this happened to the *New York Times.*[3] It was discovered that over a period of three months, 45 items of software were installed in the New York Times computer system. The New York Times relied on a Symantec anti-virus product, which only found only one item of the malicious software. Symantec subsequently posted the following comment in connection with this allegation:[4]

> Advanced attacks like the ones the New York Times described in the following article [N. Perlroth, 'Hackers in China attacked The Times for last 4 months', *New York Times*, 30 January 2013], underscore how important it is for companies, countries and consumers to make sure they are using the full capability of security solutions. The advanced capabilities in our endpoint offerings, including our unique reputation-based technology and behaviour-based blocking, specifically target sophisticated attacks. Turning on only the signature-based anti-virus components of endpoint solutions alone are not enough in a world that is changing daily from attacks and threats. We encourage customers to be very aggressive in deploying solutions that offer a combined approach to security. Anti-virus software alone is not enough.

1 J A P Marpaung, M Sain and Hoon-Jae Lee, 'Survey on malware evasion techniques: state of the art and challenges', *Advanced Communication Technology (ICACT), 2012 14th International Conference* (Global IT Research Institute 2012), pp. 744–9.

2 Daniel Bilar, 'Known knowns, known unknowns and unknown unknowns: anti-virus issues, malicious software and internet attacks for non-technical audiences' (2009) 6 Digital Evidence and Electronic Signature Law Review 123; in 2006, Graham Ingram, the general manager of the Australian Computer Emergency Response Team (AusCERT), told an audience in Sydney, Australia, that popular desktop antivirus applications do not work, reported by Munir Kotadia, 'Eighty percent of new malware defeats antivirus' (ZDNet Australia, 19 June 2006); Michael A Caloyannides, 'Digital evidence and reasonable doubt' (2003) 1 IEEE Security and Privacy 89; Dmitry Silnov, 'Features of virus detection mechanism in Microsoft Security Essentials (Microsoft Forefront Endpoint Protection)', (2013) 4 Journal of Information Security 124; also see the annual 'X-Force Trend Statistics' by IBM Internet Security Systems that reinforces the position on the failure of anti-virus software, available online at <www-03.ibm.com/security/xforce/downloads.html>; the reports produced by the Anti-Phishing Working Group (<www.antiphishing.org>) illustrate the same problem; reports by AV-Comparatives.org appear to indicate that some of the best products are now very efficient, available at <www.av-comparatives.org>; see also 'Common vulnerabilities and exposures', available at <https://cve.mitre.org>.

3 Nicole Perlroth, 'Hackers in China attacked The Times for last 4 months', *New York Times* (New York, 30 January 2013); in 2014, it was accepted by Symantec that anti-virus was no longer to be relied upon, for which see Danny Yadron, 'Symantec develops new attack on cyberhacking declaring antivirus software dead, firm turns to minimizing damage from breaches', *Wall Street Journal* (New York, 4 May 2014).

4 <www.symantec.com/connect/blogs/symantec-statement-regarding-new-york-times-cyber-attack>.

6.102 It is a truth universally acknowledged that the majority of hackers concentrate on the most widely used software and on vulnerable applications that can be found by using Internet search engines, although the development of the Stuxnet virus illustrates that governments are now probably responsible for some of the most effective viruses that are written, although organized criminals can be equally effective.[1] Software

need only include a low number of defects to create enough vulnerabilities for serious hackers to manipulate the defects to their advantage. Jim Nindel-Edwards and Gerhard Steinke usefully sum up the position:

> It would seem that after decades of software development there would be some assurance that software works as specified in the customer requirements. Is it that software vendors are unwilling to perform sufficient testing? Is it possible to test everything? Finding a certain number of bugs, doesn't mean that the software has no more bugs. On the other hand, not finding any defects doesn't mean there aren't any defects in the software either. Perhaps there are known bugs, but the time and resources to fix these bugs and defects are often not provided and the software is released with known (but not publicly stated) bugs. Is it because there is a low expectation of quality? Is it even possible to get rid of all bugs, especially when we are integrating components from multiple sources and we are dependent on the software that was developed and tested by others?
>
> Software quality assurance is a challenging task. There are many questions raised by software being released with defects. What are the ethical responsibilities of a software vendor releasing software with bugs, especially if it is system-critical software, but also when releasing non system-critical software?

1 Roderic Broadhurst, Peter Grabosky, Mamoun Alazab, Brigitte Bouhours and Steve Chon, 'Organizations and cyber crime: an analysis of the nature of groups engaged in cyber crime', (2014) 8 International Journal of Cyber Criminology 1, available at <www.cybercrimejournal.com/broadhurstetalijcc2014vol8issue1.pdf>.

2 Jim Nindel-Edwards and Gerhard Steinke, 'Ethical issues in the software quality assurance function' (2008) 8 Communications of the IIMA 53, 54.

Software testing

6.103 Most software organizations test their products extensively, including in the ways that they anticipate that their customers will use them. Indeed, most software has become so complex that in a process called 'beta testing', software has been provided to volunteers to test before it is sold as a product. It has also been suggested that the problems of the composition of components in large systems can be mitigated by programmers reusing components in ways that they know from experience tend to work,[1] although this view is not generally accepted.[2] However, there will continue to be malfunctions, because many problems in hardware, software and configuration are only exposed when the system runs under real workloads.[3] A number of issues arise in this respect, including the use of tools to test software fault tolerance or robustness,[4] the degree to which the testing accurately reflects the way users will actually use the software, how people may attempt to use the product in an unconventional way, and testing how the software works when connecting and communicating with different software and hardware. It is well known that testing software is inadequate to uncover errors, because there is never enough time to cover all the cases, as the illustrations mentioned in this chapter vividly shows. Professor Thimbleby has indicated that the only solutions are:

> (i) a very careful approach to reasoning about the requirements that lead to the decisions,
>
> (ii) a mathematically rigorous way to analyse the combinations of decisions,
>
> (iii) rigorous testing, primarily to uncover whether there were flaws in steps (i) and (ii), including in the testing process itself, and

> (iv) external oversight to avoid mistakes in one's reasoning – this includes processes such as code review by third parties.[5]

1 C A R Hoare, 'How did software get so reliable without proof?' in Marie-Claude Gaudel and Jim Woodcock (eds), *Lecture Notes in Computer Science, vol 1051/1996* (Springer 1996) 1–17.
2 Bev Littlewood and Lorenzo Strigini, 'The risks of software', *Scientific American*, 267 (November 1992) 62–75, cited by Partridge, *The Seductive Computer*, p. 205, fn. 15; Bev Littlewood and Lorenzo Strigini, 'Validation of ultra-high dependability – 20 years on' (2011) 20 SCSC Newsletter <www.staff.city.ac.uk/~sm377/ls.papers/2011_limits_20yearsOn_SCSC/BL-LS-SCSSnewsletter2011_02_v04distrib.pdf>.
3 Schroeder and Gibson, 'A large-scale study of failures in high-performance computing systems', 343.
4 Although the availability of such tools does not mean that developers use such tools to improve their systems, for which see John DeVale and Philip Koopman, 'Robust software – no more excuses' in Danielle C Martin (ed.), *Proceedings International Conference on Dependable Systems and Networks* (The Institute of Electrical and Electronics Engineers, Inc. 2002) 145–54.
5 Personal communication with the author.

6.104 The problem with a presumption that a computer is deemed to be 'reliable' is that as systems become more complex, it has become progressively more challenging to test software to reflect the way the users will actually use the product. This is because of the large number of functions that software is required to perform, and the unpredictability of the users.[1] Professor Partridge reiterates the point that 'no significant computer program is completely understood',[2] and goes further by indicating that systems are now so complex that humans are no longer able to deal with the problems:

> We might speculate further: if the nature of computer-system complexity really is new and peculiar, a system characteristic that has no parallel in the natural world, then our evolutionary history is unlikely to have equipped us to reason effectively with such systems. Our genetic programs may be totally lacking in mechanisms that can deal effectively with discrete complexity.[3]

1 The rise in fraud that took advantage of the faults in software was rapidly increasing in the 1970s, for which see Linden, 'Operating system structures to support security and reliable software' 410.
2 Derek Partridge, *What makes you clever – the puzzle of intelligence* (World Scientific 2014) 394 and 407 fn. 22.
3 Partridge, *The Seductive Computer*, 192.

6.105 This weakness is now recognized by some of the organizations that produce devices and software. Microsoft and Apple are among a number of companies that have adopted a 'bug' bounty programme to reward professionals who test and find errors in the software.[1] The US Department of Defense has also taken this approach, as has Google in respect of cryptographic software libraries.[2] Yet claims that software code and hardware products have been independently tested does not necessarily lead to the conclusion that they can be relied upon. In his ACM Turing Lecture of 1972,[3] Professor Dijkstra said this of testing:

> Today a usual technique is to make a program and then to test it. But: program testing can be a very effective way to show the presence of bugs, but is hopelessly inadequate for showing their absence.

1 Microsoft Bounty Programs <https://technet.microsoft.com/en-us/library/dn425036.aspx>; Hannah Kuchler, 'Apple offers $200,000 bounty to identify flaws in its software', *Financial Times Companies & Markets*, (6/7 August 2016), p. 11.
2 DoD Vulnerability Disclosure Policy, available at <https://hackerone.com/deptofdefense>; Project

Wycheproof (19 December 2016) <https://security.googleblog.com/2016/12/project-wycheproof.html>.

3 Available at <www.cs.utexas.edu/~EWD/ewd03xx/EWD340.PDF>; the lecture was published as an article: E W Dijkstra, 'The humble programmer' (1972) 15 Communications of the ACM 859.

6.106 In other words, good quality testing might discover the failings of the developer, but are less capable of resolving the issues in the overall design of software: there are significant limits to testing.

Writing software that is free of faults

6.107 As Professor Thomas indicates, it is possible to design and develop software so that it is almost completely free of faults.[1] Many applications are now built without the developer writing any code at all. The coding is done in building the tools that generate the code when given parameters by the developer – and this is premised on the fact that the software tools that generate the code are themselves error free.

1 'Should we trust computers?', a lecture given at Gresham College on 20 October 2015, available at <www.gresham.ac.uk/lectures-and-events/should-we-trust-computers>.

Software standards

6.108 Where an organization produces safety critical software for airplanes, motor vehicles, air traffic control, or power stations, it will be necessary to conform to the requirements of an international standard on functional safety of programmable electronic systems.[1] For instance, security in the banking sector relies on certification standards such as FIPS-140 Information Technology Security Evaluation Criteria (ITSEC) and the Common Criteria for Information Technology Security Evaluation. It should be noted that these schemes only focus on aspects of security, and not on overall functionality. It is possible to have an accredited product that implements the security functions well, but its business functions badly.

1 For a discussion, see the NuSAC Study Group on the Safety of Operational Computer Systems, *The use of computers in safety-critical applications* (HMSO 1998).

6.109 The ITSEC scheme, which is no longer as active as it once was, assesses an organization's product based on document prepared by the organization that wants that product to be evaluated. In general terms, a document that is submitted to ITSEC describes what the product is designed to do, the situation in which it is intended to operate in, the risks the product is likely to encounter, and the mechanism by which the product acts to protect against the risks. It is for ITSEC to determine whether the claims are substantiated. Only the risks identified by the applicant are tested. A product is given one of seven levels from E0 (no formal assurance) to E6 (the highest level of confidence). Each level represents increasing levels of confidence. The assessment and granting of a position on the E scale is a judgment that a certain level of confidence has been met. It is not a measure of the strength of the security in place. It is important to realize that the organization submitting the product for evaluation sets out the criteria by which it will be evaluated. It may be that the party submitting the product for evaluation will not have included the risks associated with the use of the product by the end user. The evaluation includes an assessment of the confidence to be placed in whether the security features are the correct ones and how effective the security

features work. This means that a security mechanism might be applied correctly, but it will not be effective unless it is appropriate for the purpose for which they have been designed. In this respect, it is necessary to know why a particular security function is necessary, what security is actually in place, and how the security is provided. It does not follow that if a product has a high E level, it will provide a high level of security.

6.110 The 'Common Criteria for Information Technology Security Evaluation' and 'Common Methodology for Information Security Evaluation' comprise the technical basis for an international agreement called the 'Common Criteria Recognition Agreement'. The manufacturer submits its product to an independent licensed laboratory for an assessment of the product. The way a product is evaluated is similar to the way ITSEC undertakes such assessments. There are problems with this, because it creates a conflict of interest: there are no known examples of the revocation of licences of laboratories that conduct evaluations, both parties are able to subvert the process, and determining the name of the organization that conducted the evaluation might be impossible without an order for disclosure. In addition, claims will sometimes be made that a device has been certified when, in fact, it might only have been evaluated. Often, a bank will ask a judge to rely on the certification process without disclosing the relevant report. For instance, it has been demonstrated that independent external examination continues to validate and approve of devices and cryptographic software code that are open to failure and subversion.[1]

1 Steven J Murdoch, Mike Bond and Ross Anderson, 'How certification systems fail: lessons from the Ware Report' (2012) 10 IEEE Security & Privacy 40; Kim Zetter, 'In legal first, data-breach suit targets auditor', *Wired* (16 February 2009) – the case mentioned in this article was *Merrick Bank Corporation v Savvis, Inc.*, 2010 WL 148201 (for other references, see 2009 WL 2968844 (D.Ariz.) (Trial Motion, Memorandum and Affidavit) (5 June 2009); 2009 WL 4823623 (D.Ariz.) (Trial Motion, Memorandum and Affidavit) (7 July 2009); 2009 WL 4823624 (D.Ariz.) (Trial Motion, Memorandum and Affidavit) – it is not clear what happened as a result of the legal action. It is probable that the case was settled after the court refused to dismiss the case. Another case, a class action, was initiated in the United States District Court Northern District of Illinois Eastern Division on 24 March 2014: *Trustmark National Bank v Target Corporation*, Case NO 14-CV-2069, although it was reported that this action was subsequently withdrawn, for which see J Stempel, 'Banks pull out of lawsuit vs Target, Trustwave over data breach', *Reuters* (1 April 2014).

6.111 Two observations are worthy of note: that standards (the use of standards is a topic of significant debate, because it is not always certain that they work to improve the quality of software) regarding aviation, space and medical devices are usually much more prescriptive that those used in other domains, and even within the aviation, space and medical industries, a great deal of commercial software is developed against no formal process model at all. The relevant standard for medical devices is 'ISO 13485:2003 Medical devices – Quality management systems – Requirements for regulatory purposes' (now revised by 'ISO 13485:2016 Medical devices – Quality management systems – Requirements for regulatory purposes'). This standard has historically placed much less focus on tracing the details of internal product structure than, for instance, DO-178B, Software Considerations in Airborne Systems and Equipment Certification, which is a guideline dealing with the safety of safety critical software to be used in certain airborne systems. Yet, although having software evaluated against standards is a laudable goal, it does not follow that by conforming, errors are eliminated.[2]

1 By way of example, see Patrick J Graydon and C Michael Holloway, 'Planning the unplanned experiment: assessing the efficacy of standards for safety critical software' (NASA/TM{2015{218804, September 2015), at <http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20150018918.pdf>.
2 Timothy J Shimeall and Nancy G Leveson, 'An empirical comparison of software fault tolerance and fault elimination' (1991) 17 Transactions on Software Engineering 173; P B Ladkin, 'Opinion – taking software seriously' (2005) 41 Journal of System Safety, available at <www.rvs.uni-bielefeld.de/publications/Reports/Ladkin-JSS-May2005.pdf>; Harold Thimbleby, Alexis Lewis and John Williams, 'Making healthcare safer by understanding, designing and buying better IT' (2015) 15 Clinical Medicine 258.

Summary

6.112 In summary, faults in software and errors relating to the design of software systems are exceedingly common.[1] And while defects in hardware have been relatively rare,[2] they are not unknown.[3] Hardware is increasingly developed using high-level languages similar to those used for software. Furthermore, hardware is being released with firmware which may be reconfigured for other purposes. In addition, hardware faults can also be introduced by the improper use or configuration of software tools designed for developing hardware, which may themselves be error-prone. Like software, hardware errors, too, can be exploited to cause security failures.[4]

1 L Strigini, 'Fault tolerance against design faults', in Hassan B Diab and Albert Y Zomaya (eds.), *Dependable Computing Systems: Paradigms, Performance Issues, and Applications* (Wiley 2005), 213–41.
2 Such as the Pentium FDIV or 'floating point' error (although this, strictly speaking, was a software fault), although Intel could not fix the error other than to issue a replacement. Professor Thomas R. Nicely was the first to publicize this fault, Partridge, *The Seductive Computer*, 98, fn. 8. For further information, see <www.trnicely.net/#PENT> and <www.cs.earlham.edu/~dusko/cs63/fdiv.html>; see also the FDIV Replacement Program – White Paper: Statistical Analysis of Floating Point Flaw (30 November 1994), Intel Corporation (this is no longer available online). Note also a recent paper by Bianca Schroeder, Eduardo Pinheiro and Wolf-Dietrich Weber, 'DRAM errors in the wild: A large-scale field study' in which the incidence of memory errors and the range of error rates across different DIMMs (dual in-line memory modules) are much higher than previously reported, for which see <www.cs.toronto.edu/~bianca/papers/sigmetrics09.pdf>; Bianca Schroeder and Garth A Gibson, *Disk failures in the real world: What does an MTTF of 1,000,000 hours mean to you?* (Proceedings of the 5th USENIX Conference on File and Storage Technologies (FAST'07), February 2007), 2006 version available at <www.pdl.cmu.edu/PDL-FTP/Failure/CMU-PDL-06-111.pdf>; Eduardo Pinheiro, Wolf-Dietrich Weber and Luiz Andre Barroso, *Failure Trends in a Large Disk Drive Population* (Proceedings of the 5th USENIX Conference on File and Storage Technologies (FAST'07), February 2007), available at <http://static.googleusercontent.com/media/research.google.com/en//archive/disk_failures.pdf>.
3 Most complex integrated circuits in wide use will have published lists of 'errata' – for example, Intel published 'Intel core duo processor and Intel core solo processor on 65 nm process specification update June 2009 revision 020', which lists 84 separate items, available at <http://download.intel.com/design/mobile/SPECUPDT/30922214.pdf>.
4 For an example, see Sudhakar Govindavajhala and Andrew W Appel 'Using memory errors to attack a virtual machine', available at <www.cs.princeton.edu/~appel/papers/memerr.pdf>.

6.113 Every part of a program is different, and must be made correct independently. In the case of machines, there are two important differences: things are almost always continuous, and after a time, the system is back where it started. When they are not continuous, problems always occur. For example, a wheel turns, and once it has turned, it is (not withstanding wear and tear) likely to be able to turn again. Each time it turns, it gets back to an indistinguishable state. This is called a symmetry. Symmetries are very general ideas. For example, if one moves a cup of coffee a foot to the left, it

stays the same, and works exactly as before. This is because the world we live in has translational symmetry – everything is the same if it is moved. Wheels have rotational symmetry, and so on. This means that almost all of the design decisions in mechanical devices 'collapse' because of symmetries, and there is not the exponential growth of cases that happens in software. On the other hand, no part of a software program is the same as any other part. Indeed, if it was, one would ask why it was so inefficiently designed. Thus there are no symmetries in software that amplify the 'how it works' thinking that so readily simplifies physical design.

6.114 In particular, it might be obvious that the behaviour of a stopwatch used by a policeman is the 'same' as the behaviour of the 'same' stopwatch presented in court as evidence, or in the laboratory where it was tested. Thanks to symmetries, moving a watch from the roadside to the laboratory does not change it. There is no symmetry to justify software adduced in court behaving as it did anywhere else. Software is not constrained, as any physical device is, to work in the universe with all its symmetries. Software does not obey any of them, and thanks to human error (known and unknown) in its design, its behaviour cannot be taken for granted.[1]

1 I owe this analysis to Professor Thimbleby.

6.115 Software will continue to be unreliable. By providing a general presumption of reliability to software, the law acts to reinforce the attitude of the software industry that the effects of poor quality work remain the problem of the end user. In many circumstances, because the user can himself cause errors, the industry may seek to pin the blame on the user himself, further obfuscating the true origin and source of the errors.[1] For these reasons, it is rare for a customer to take legal action against the software supplier, let alone attempt such an action, and be successful.[2]

1 The various pressures are illustrated in David Hechler, 'Lost in Translation?' (April 2013) Corporate Counsel 72, available at <www.asbpe.org/blog/2014/07/28/david-hechler-wins-asbpes-2014-stephen-barr-award-for-article-on-toyotas-fatal-acceleration-problems/>. David Hechler is the executive editor of *Corporate Counsel* magazine, and the American Society of Business Publication Editors gave him the 2014 Stephen Barr Award for this article.

2 For example, see the English cases of *St Albans City and District Council v International Computers Limited* [1996] 4 All ER 481, [1997] FSR 251 and *Kingsway Hall Hotel Ltd v Red Sky IT (Hounslow) Ltd* [2010] EWHC 965 (TCC). Elizabeth MacDonald considered the position in contract, giving a number of examples in her article 'Bugs and breaches' (2005) 13 Intl J L & Info Tech 118. National Air Traffic Services initiated action against Electronic Data Systems Ltd, although the outcome is not certain. For an appeal against an application to amend the reply and defence to counterclaim, see *Electronic Data Systems Ltd v National Air Traffic Services* [2002] EWCA Civ 13 – Professor Ladkin indicated that the software development could fail, for which see Memorandum by Professor Peter B Ladkin (ATC 20) submitted to the Select Committee on Environment, Transport and Regional Affairs Fourth Report (ordered by the House of Commons to be printed 27 March 1998), available at <www.publications.parliament.uk/pa/cm199798/cmselect/cmenvtra/360-e/36082.htm>.

6.116 This discussion apart, the central issue for lawyers is dealing with the presumption that a computer is working properly. The following summary of the problems of software by Professor Partridge help to remind us of the landscape:

> IT systems are everywhere, and will continue to infiltrate the lives of all of us.
>
> We cannot easily check that an IT system is computing correctly.
>
> IT systems all fail: sometimes immediately and spectacularly, sometimes unobtrusively just once in a while, and sometimes in any combination of these

> two extremes.
>
> IT-system failures vary from production of blatantly incorrect results to failure to produce a desired result.
>
> The interplay of a variety of causes means that all large IT systems are unmanageably complex.
>
> IT-system complexity is discrete complexity rather than complexity based on continua.
>
> If, by chance (combined with exemplary practice and much effort), an IT system is constructed with no possibility of failure behaviour, we can never know this.[1]

1 Partridge, *The Seductive Computer*, 9.

6.117 This poses a question for lawyers, experts and the courts: how should the reliability of software be reviewed in a court of law?

Challenging 'reliability'

6.118 When seeking to challenge the underlying software of a computer or computer-like device, lawyers frequently have great difficulty in overcoming the presumption that a machine is working properly, although general assertions about the failure of software code are often made without providing any foundation for the allegations. This problem is compounded when a party refuses to deliver up relevant evidence, usually citing confidentiality as the reason for the refusal, and relying on the presumption that a computer is 'reliable'. In such circumstances, it is difficult to convince a judge to order the disclosure of relevant data.

6.119 Yet, paradoxically, it is a well-known fact in the industry that software could hardly be said to be 'reliable'. As noted by Steyn J in *Eurodynamic Systems Plc v General Automation Ltd*:

> … The expert evidence convincingly showed that it is regarded as acceptable practice to supply computer programmes (including system software) that contain errors and bugs. The basis of the practice is that, pursuant to his support obligation (free or chargeable as the case may be), the supplier will correct errors and bugs that prevent the product from being properly used.[1]

1 (6 September 1988, not reported), QBD, 1983 D 2804, [5.a].

6.120 This view is reinforced by Professor Matt Blaze:

> It is a regrettable (and yet time-tested) paradox that our digital systems have largely become more vulnerable over time, even as almost every other aspect of the technology has (often wildly) improved.
>
> …
>
> Modern digital systems are so vulnerable for a simple reason: computer science does not yet know how to build complex, large-scale software that has reliably correct behaviour.[1] This problem has been known, and has been a central focus of computing research, since the dawn of programmable computing. As new technology allows us to build larger and more complex systems (and to connect them together over the internet), the problem of software correctness becomes exponentially more difficult.[2] [Footnote 2 is at this point, and is reproduced below]

Footnote 2:

> That is, the number of software defects in a system typically increases at a rate far greater than the amount of code added to it. So adding new features to a system that makes it twice as large generally has the effect of making far more than twice as vulnerable. This is because each new software component or feature operates not just in isolation, but potentially interacts with everything else in the system, sometimes in unexpected ways that can be exploited. Therefore, smaller and simpler systems are almost always more secure and reliable, and best practices in security favour systems [that have] the most limited functionality possible.[3]

1 It should be noted that computer scientists have invented many ways to achieve this, and some companies use these methods to prove mathematically that their systems cannot fail at runtime – but the software will be running on a computer with unreliable hardware, other firmware and software and user interfaces, which might mean that the program might be 'right', but when interacting with the other components, can lead to a lethal failure. Also, we need to be aware that what is being proved is not that the systems do what is desired, but that the systems meet a formal statement of the requirements. The original requirements cannot be themselves be proved to be correct, or that the formal software requirements meet the constraints of the real world. There are limits to what formal methods can do, and those limits are not widely acknowledged. See B Littlewood and L Strigini, 'Validation of ultrahigh dependability for software-based systems' (1993) 36 Communications of the ACM 69, available at <http://openaccess.city.ac.uk/1251/1/CACMnov93.pdf>.
2 It is not clear whether 'exponentially' means that the rate of growth is proportional to the amount present, or whether the word is used loosely to mean 'growing rapidly'.
3 Matt Blaze, Testimony to the Subcommittee on Information Technology hearing, 'Encryption Technology and Potential U.S. Policy Responses' on Wednesday, April 29, 2015 at 2:00pm, available at <http://democrats.oversight.house.gov/legislation/hearings/subcommittee-on-information-technology-hearing-encryption-technology-and>.

6.121 Lawrence Bernstein and C. M. Yuhas also acknowledged this observation:[1]

> Software developers know that their systems can exhibit unexpected, strange behaviour, including crashes or hangs, when small operational differences are introduced.[2] These may be the result of new data, execution of code in new sequences or exhaustion of some computer resource such as buffer space, memory, hash function overflow space or processor time.

1 Lawrence Bernstein and C M Yuhas, 'Design constraints that make software trustworthy' *IEEE Reliability Society 2008 Annual Technology Report* 3, available at <http://rs.ieee.org/tech-activities/42-letters-in-reliability-annual-technology-reports>.
2 This is a consequence of discrete complexity, or digital complexity.

6.122 This section aims to provide a broad outline of the problems relating to computers and computer-like devices by different industries, and to illustrate the importance of software and how there may be times when the output of a computer may not necessarily be 'reliable' and is therefore not to be trusted. Software code should be open to scrutiny, and should not necessarily share the benefit of a presumption of 'reliability' that is incapable of being effectively challenged.[1]

1 Ken Chasse, 'Electronic records as documentary evidence' (2007) 6 Canadian Journal of Law and Technology 141 refers to the need for a 'system integrity test'.

6.123 One of the problems with understanding the role of the presumption is that people fail to distinguish software from computer systems. Computers are merely devices that are remarkable in that they can be turned to do many tasks rather than being limited to a single purpose. In order to perform a useful purpose, they must be

instructed by software. A computer and its software together can be taken to form a *system*. No machine is 'reliable' or 'unreliable' in an absolute sense. Machines may be *more* or *less* reliable. The term 'reliable' in everyday use is an abbreviation of what in technical terms is 'reliable enough for the intended purpose'. All machines have some probability of failing, so none is 'reliable' in the sense that one can rely on it without any doubt, while many are reliable enough (their probability of failing to perform correctly at any one use is small enough) to be worth using. The problem with using the word 'reliable' as though reliability were a binary quality, is that we risk taking it to mean 'reliable enough' without allowing of the fact that what is 'enough' depends on the use to which we put the machine, or rather, its outputs. For instance, a machine may be reliable enough to be worthwhile in everyday use, and yet not reliable enough to use as evidence in a specific case. The speedometer in a motor car may be reliable enough to use as an aid for driving at reasonable speed, because this level of reliability is not necessarily the same level of reliability that should be required in order to use not a matter of whether the instrument is 'reliable', but of 'how reliable' it is. It follows that lay people are not aware of the inherent design faults, and trust their personal experience to reassure themselves that computers are 'reliable' machines. Yet lay users experience problems with devices regularly, which illustrates the failure of lay people to grasp that 'reliability' and software code are impossible to guarantee.[1]

1 David Harel, *Computers Ltd. What They Really Can't Do* (Oxford University Press 2003); see also Neumann, *Computer Related Risks*, and his website, which is continually updated: <www.csl.sri.com/users/neumann/insiderisks.html>; see also the list of software failures on the website of Nachum Dershowitz, School of Computer Science, Tel Aviv University, at <www.cs.tau.ac.il/~nachumd/horror.html>.

6.124 Lay people are not the only people to make this mistake. This may be illustrated by the judicial assertion that computers are 'reliable' because there are more computers. Villanueva JAD made just such an assertion without providing any evidence to sustain his claim that computers are 'presumed reliable' in the case of *Hahnemann University Hospital v Dudnick*:

> Clearly, the climate of the use of computers in the mid-1990's is substantially different from that of the 1970's. In the 1970's, computers were relatively new, were not universally used and had no established standard of reliability. Now, computers are universally used and accepted, have become part of everyday life and work and are presumed reliable.

1 292 N.J.Super. 11, 678 A.2d 266 (N.J.Super.A.D. 1996), 268.

6.125 This observation by Villanueva JAD was made in the same year as the failure of the software that caused the *Ariane 5* rocket to be destroyed shortly after take-off.

6.126 That computers are deemed to be 'reliable' because they are used more frequently is a poor substitute for a rigorous understanding of the nature of computers and their software. However, it is accepted that long-term use can be an important element of justified trust in a software system. This comes about because there might be a long history of valuable and seemingly error-free use, but also because the long-term user typically gets to know the idiosyncrasies of the system.

Aviation industry

6.127 Errors in aviation software can have disastrous, or near disastrous, consequences. It can be caused by something as simple as bad coding. By way of example, consider the F-22A Raptor advanced tactical fighter, which entered service with the US Air Force in 2005. In February 2007, 12 of these aircraft were flying from Hickham AFB in Hawaii to Kadena AB on Okinawa. All of the aircraft experienced simultaneous and total software failure with their navigational console when their longitude shifted from 180 degrees West to 180 East. The jets were accompanied by tanker planes, which meant the pilots in the tankers were able to guide the jets back to Hawaii. Major General Don Sheppard spoke about the problem on CNN on 24 February 2007. The related part of the transcript is set out below:

> Maj. Gen. Don Sheppard (ret.): … At the international date line, whoops, all systems dumped and when I say all systems, I mean all systems, their navigation, part of their communications, their fuel systems. They were – they could have been in real trouble. They were with their tankers. The tankers – they tried to reset their systems, couldn't get them reset. The tankers brought them back to Hawaii. This could have been real serious. It certainly could have been real serious if the weather had been bad. It turned out OK. It was fixed in 48 hours. It was a computer glitch in the millions of lines of code, somebody made an error in a couple lines of the code and everything goes.
>
> [snip]
>
> SHEPPERD: Absolutely. When you think of airplanes from the old days, with cables and that type of thing and direct connections between the sticks and the yolks and the controls, not that way anymore. Everything is by computer. When your computers go, your airplanes go. You have multiple systems. When they all dump at the same time, you can be in real trouble. Luckily this turned out OK.
>
> John Roberts, CNN anchor: What would have happened General Shepperd if these brand-new $120 million F-22s had been going into battle?
>
> SHEPPERD: You would have been in real trouble in the middle of combat. The good thing is that we found this out. Any time – before, you know, before we get into combat with an airplane like this. Any time you introduce a new airplane, you are going to find glitches and you are going to find things that go wrong. It happens in our civilian airliners. You just don't hear much about it but these things absolutely happen. And luckily this time we found out about it before combat. We got it fixed with tiger teams in about 48 hours and the airplanes were flying again, completed their deployment. But this could have been real serious in combat.
>
> ROBERTS: So basically you had these advanced air – not just superiority but air supremacy fighters that were in there, up there in the air, above the Pacific Ocean, not much more sophisticated than a little Cessna 152 only with a jet engine.
>
> SHEPPERD: You got it. They are on a 12 to 15-hour flight from Hawaii to Okinawa, but all their systems dumped. They needed help. Had they gotten separated from their tankers or had the weather been bad, they had no attitude reference. They had no communications or navigation. They would have turned around and probably could have found the Hawaiian Islands. But if the weather had been bad on approach, there could have been real trouble. Again, you get refueling from your tankers. You don't run – you don't get yourself where you run out of fuel. You always have enough fuel and refueling nine, 10, 11, 12 times on a flight like this where you can get somewhere to land. But again, attitude reference and navigation are essential as is communication. In this case all of that was affected.

It was a serious problem.[1]

1 'F-22 Squadron shot down by the International Date Line', *Defense Industry Daily* (1 March 2007), at <www.defenseindustrydaily.com/f22-squadron-shot-down-by-the-international-date-line-03087/>; L. Page, 'US Superfighter software glitch fixed', *The Register* (28 February 2007).

6.128 In practice, it means that most commercially produced software will have thousands of undetected defects.[1]

1 For software defects generally, see Frederick P Brooks, *The Mythical Man-Month: Essays on Software Engineering* (2nd edn, Addison-Wesley 1995) and a discussion by Professor Les Hatton substantiates the broad range quoted here: *Some Notes on Software Failure* (Addison-Wesley 2001). See also Jim Nindel-Edwards and Gerhard Steinke, 'Ethical issues in the Software Quality Assurance function' (2008) 8 Communications of the IIMA 53.

6.129 In conventional flight control, the flight control commands from the cockpit are conveyed mechanically through steel cables or pushrods, often servo-assisted, to hydraulic actuators which then physically move the aerodynamic control surfaces on the wings and tailplane. In 'fly-by-wire', the flight control commands are converted to electrical signals transmitted by wires to the control surface actuators (in some cases in modern fly-by-wire aircraft the actuators may also be electric). Flight control is completely intermediated by software code, and a more accurate description would now be 'fly-by-software-code'. Besides fly-by-wire, the autopilot and flight management systems of even conventionally-controlled aircraft are software-based. The more reliable and functional the autopilot and flight management systems software have become, the more pilots have relied on it, even to the detriment of their piloting skills, as demonstrated by a number of accidents and ensuing loss of life. Accidents involving aircraft can exhibit a series of anomalous pilot-system interactions, and aviation regulations, and investigators, with few exceptions tend to assign the responsibility for the results of those interactions ultimately to the pilots. This is so even in circumstances where it is clear that the software code and the system design are so faulty that a human being is not able to respond correctly – or with sufficient speed. In the case of American Airlines Flight 965 near Cali, Colombia, on 20 December 1995,[2] 151 passengers and all of the cabin crew members died in the crash. In this case, a significant error occurred, as explained by Highsmith DJ:

> American Airlines predicates its claims on Honeywell's role as supplier of the Flight Management Computer (FMC) used on Flight 965 and Jeppesen's role in furnishing the navigational database programmed into the FMC and the corresponding aviation charts. Without making any findings in this regard but simply reflecting the narrative contained in Judge Marcus' summary judgment opinion, the Court notes that, on the approach to Cali, the pilots entered 'R' into the FMC, anticipating (based on the aviation charts) that this cipher corresponded to a beacon designated as 'Rozo'. Instead, another beacon designated as 'Romeo' was activated. This resulted in a change of the aircraft's heading to the east, over the Andes mountains. When the pilots became aware of the aircraft's easterly swing, they turned back to the west, in the direction of the valley where the Cali airport is located. Sadly, since the aircraft had been descending during these directional changes, Flight 965 never made it back to the valley. It crashed into the side of a mountain.[3]

1 Bill Palmer, *Understanding Air France 447* (Print edition v1.05, 2013), 179 and *Safety Alert for Operators*, issued by the U.S. Department of Transportation, Federal Aviation Administration (SAFO 13002 1/4/13) <www.faa.gov/other_visit/aviation_industry/airline_operators/airline_safety/safo/

all_safos/media/2013/SAFO13002.pdf>; Susan Carey, 'American Airlines flight delays continue as pilot iPad app glitch is fixed', *Wall Street Journal* (29 April 2015) <www.wsj.com/articles/american-airline-flight-delays-continue-as-pilot-ipad-app-glitch-is-fixed-1430335366>; Alex Hern, 'App fail on iPad grounds "a few dozen" American Airlines flights', *The Guardian* (29 April 2015) <www.theguardian.com/technology/2015/apr/29/apple-ipad-fail-grounds-few-dozen-american-airline-flights>.

2 *In Re Air Crash Near Cali, Colombia on December 20*, 24 F.Supp.2d 1340 (1998).

3 At 1342 (footnotes omitted).

6.130 The critical importance of verifying the design of aviation software based on industry standards was noted in the *Aviation Occurrence Investigation Final Report: In-flight upset 154 km west of Learmonth*.[1] In this case, a problem with the software controlling the aeroplane was a cause of the accident. In this investigation report, the authors cited text relating to software requirements from *Software Considerations in Airborne Systems and Equipment Certification*,[2] produced by the Radio Technical Commission for Aeronautics:

> DO-178A [now DO-178C] provided high-level guidance for the generation of software requirements, the verification that the resulting design met the requirements, and validation that the requirements were adequate. It also noted that for systems that performed certain critical and essential functions:
>
> > ... it may not be possible to demonstrate an acceptably low level of software errors without the use of specific design techniques. These techniques, which may include monitoring, redundancy, functional partitioning or other concepts, will strongly influence the software development program, particularly the depth and quality of the verification and validation effort ...
> >
> > NOTE: It is appreciated that, with the current state of knowledge, the software disciplines described in this document may not, in themselves, be sufficient to ensure that the overall system safety and reliability targets have been achieved. This is particularly true for certain critical systems such as full authority fly-by-wire. In such cases it is accepted that other measures, usually within the system, in addition to a high level of software discipline may be necessary to achieve these safety objectives and demonstrate that they have been met.[3]

1 WA 7 October 2008 VH-QPA Airbus A330-303 (ATSB Transport Safety Report, AO-2008-070).

2 (DO-178A, SC-152, issued on 22 March 1985 and updated regularly) <www.rtca.org>.

3 At 2.3.5.

Financial products

6.131 In August 2006, the rating agency Moody's rated constant proportion debt obligations (CPDOs) with an AAA rating, which was close to making an investment in a CPDO free of risk.[1] In comparison, another rating agency, Fitch, a competing rating agency, could not understand why such a high rating was given to such 'investments', because its own models put CPDOs at almost the grade of 'junk'.[2] It transpired that the software used by Moody's for the purpose of rating CPDOs had a number of faults. A fault was found in early 2008 that, when corrected, failed to give the AAA rating, increasing the likelihood of defaults. The rating committee failed to disclose the error to investors or clients, and although the error was eventually corrected, other changes were made to the code to ensure the AAA rating continued to be forthcoming.[3] A subsequent external investigation by the law firm Sullivan & Cromwell established that

members of staff had engaged in conduct contrary to Moody's Code of Professional Conduct.[4] Moody's subsequently received a 'Wells Notice'[5] from the SEC on 18 March 2011.[6] The Division of Enforcement of the Securities and Exchange Commission later issued a *Report of Investigation* into the matter.[7] In a section of the *Report*, there was an examination of the attitude of the people responsible for dealing with the software error. It is revealing, and it merits setting out in full:

> B. Rating Committee Conduct
>
> MIS subsequently held several internal rating committee meetings in France and the United Kingdom to address the coding error. MIS corrected the coding error on February 12, 2007, but made no changes to the outstanding credit ratings for CPDO notes at that time. Internal e-mails show that committee members were concerned about the impact on MIS's reputation if it revealed an error in the rating model. A January 24, 2007, e-mail from a rating committee member to the Team Managing Director chairing the committee stated:
>
> > In this particular case we seem to face an important reputation risk issue. To be fully honest this latter issue is so important that I would feel inclined at this stage to minimize ratings impact and accept unstressed parameters that are within possible ranges rather than even allow for the possibility of a hint that the model has a bug.
>
> On April 27, 2007, after additional analysis, the rating committee voted not to downgrade the affected credit ratings for the CPDO notes. The committee members felt that because the CPDO notes were generally performing well there would be no ostensible justification for downgrading the credit ratings, absent announcing the coding error. In declining to downgrade the credit ratings, the committee considered the following inappropriate non-credit related factors: (i) that downgrades could negatively affect Moody's reputation in light of ongoing negative media focus in Europe on Moody's Joint Default Analysis; (ii) that downgrades could impact investors who relied on the original ratings; and (iii) the desire not to validate the criticisms of Moody's ratings of CPDOs that had been made by a competitor and covered in the local media. The committee was comprised of senior level staff, including two Team Managing Directors, two Vice President-Senior Credit Officers, and a Vice President-Senior Analyst.

1 For the broader picture, see Charles W Calomiris and Stephen H Haber, *Fragile by Design: The Political Origins of Banking Crisis and Scarce Credit* (Princeton University Press 2014), 266–9.

2 The same scepticism was expressed by Richard Beales, Saskia Scholtes and Gillian Tett with Paul J Davies, 'Failing grades? Why regulators fear credit rating agencies may be out of their depth', *Financial Times* (London, 17 May 2007), 13.

3 This was revealed by Sam Jones, Gillian Tett and Paul J Davies, 'Moody's error gave top ratings to debt products', *Financial Times* (London, 20 May 2008).

4 S. Jones, 'When junk was gold', *FT Weekend* (London, 18/19 October 2008), pp. 16–22.

5 A 'Wells Notice' is a letter sent by a securities regulator to a prospective respondent, notifying him of the substance of charges that the regulator intends to bring against the respondent, and affording the respondent with the opportunity to submit a written statement to the ultimate decision maker.

6 Phil Wahba, 'UPDATE 2-Moody's says got Wells Notice from SEC', *Reuters* (7 May 2010).

7 Release No. 62802/31 August 2012, available at <www.sec.gov/litigation/investreport/34-62802.htm>.

6.132 Because the rating committee met in France and the UK and not in the US, the SEC declined to take any further action, '[b]ecause of uncertainty regarding a jurisdictional nexus to the United States in this matter'.

6.133 Although the SEC declined to take action in this case, it did take action against AXA Rosenberg Group LLC, AXA Rosenberg Investment Management LLC and Barr Rosenberg Research Center LLC. In this instance, an employee discovered an error in the computer code of a quantitative investment model used to manage client portfolios. The employee brought the matter to the attention of senior management. The employee was told to keep quiet about the error and not to inform others about it. The error adversely affected 608 of 1,421 client portfolios managed by AXA Rosenberg Investment Management and caused US$216,806,864 in losses. Cease-and-desist proceedings were instituted and the respondents were jointly and severally ordered to pay a civil money penalty in the amount of US$25m to the US Treasury.[1]

1 The order is available at <www.sec.gov/litigation/admin/2011/33-9181.pdf>.

6.134 Another example that might be considered to be mundane is that of software systems for the use of stockbrokers. Stockbrokers used to be regulated by the Financial Services Authority (FSA) (they are now regulated by the Financial Conduct Authority), and were required to conduct their business in accordance with relevant legislation and the rules laid out by the FSA. Failure to follow the rules may cause the FSA to take disciplinary action against the firm. In the case of *SAM Business Systems Limited v Hedley and Company (sued as a firm)*,[1] the partners of Hedley used to handle their stockbroking business with a system known as ANTAR, but late in 1999 they decided it might not work after the century date change, so they decided to buy a new product from SAM, a small software company whose only product was an item of software known as InterSet. SAM claimed this product was a ready-made package of software modules made by SAM for stockbrokers and others (such as banks) dealing in stocks and shares in administering their systems. Hedley agreed to buy the new system, but immediately after the system went live, serious problems were apparent, many of which were fixed, some speedily. (The word 'fix' is the telling word here: a local fix within a large and complex piece of software often generates problems elsewhere). Hedley continued to use InterSet, but problems persisted. Eventually, they decided to find another product for their purposes. In his judgment, Judge Bowsher QC discussed the issue of defaults in software:

> The point has frequently been made during the trial that InterSet works well elsewhere (and I have received evidence from stockbrokers, Hoodless Brennan to that effect) and accordingly it is said, if it did not work for Hedley's there must be something wrong with Hedley's method of working. That line of argument has prompted me to ask, (a) if it is a tried and tested system, why when supplied to Hedley's did it have admitted bugs? (b) what is the difference between a bug and a defect?[2]

1 [2002] EWHC 2733 (TCC), [2003] 1 All ER (Comm) 465.
2 [2002] EWHC 2733 (TCC) at [19], [2003] 1 All ER (Comm) 465, [20].

6.135 The full nature of the problems encountered with this software that purported to be written for the specific purpose for which it was supplied merits setting out in full:

> To complete the history, I must mention a document produced at my request as Exhibit C2. During the evidence of Mr. Whitehouse, I asked for a copy of a timesheet to which he had referred. That is a timesheet of 'maintenance activity' for which no charge was made. That document had not been disclosed until I

> asked for it. It is a document of 10 pages. I have not counted each item, but there are about 35 items on each of the first 9 pages and 16 on the last page. According to the claimants, the hours worked amount to 785.25. The period of time covered by the document is from 4 January 2000 to 7 February, 2001. The majority of those items appear to be efforts to fix defects. The fact that no charge was made suggests that all items fall into that category. I am not going to go through all of that document, but I will take one example. On 12 January, 2001, there is an entry, 'Analysing the problems with Hedley contract report … problem actually with contract form not the report'. On 15 January a temporary fix was prepared. On 15, 16 and 17 January over 17 hours are recorded working on this problem. Then on 17 January there is another entry, "Attempting to find the reason for the intermittent bad contracts. Not found yet". On 18 January, 2001, there is an entry, 'Attempting to find the reason for the intermittent bad contracts. The reason appears to be conflicting requirements of procedures. Needs deeper understanding of form'. There were then further entries for modifications to put the problem right on 19, 23, 24, 25 and 26 January, 2001. More work was done on the same problem on 5, 7, and 9 February, 2001. On 5 February, 2001, changes were made, 'To prevent contracts being saved where the values do not add up'. Through February, 2001 there was a series of calls to deal with a problem with split deals commission. In mid April, 2001 there was a problem with trial balances. It is quite clear from that document, produced only under pressure during the trial, as well as from all the other evidence to which I have referred, that InterSet as delivered to Hedley's was never in satisfactory working order.[1]

1 [2002] EWHC 2733 (TCC) at [128], [2003] 1 All ER (Comm) 465, [129].

6.136 Two experts were appointed to give evidence in this case, and they signed an agreement which was, is in fact, a schedule of defects alleged by Hedley with comments on each from SAM. This schedule of faults ran to 34 pages. Judge Bowsher QC offered some pertinent comments in relation to the attitude of the software supplier in this case:

> SAM, like some others in the computer industry seem to be set in the mindset that when there is a 'bug' the customer must pay for putting it right. Bugs in computer programmes are still inevitable, but they are defects and it is the supplier who has the responsibility for putting them right at the supplier's expense.[1]

1 [2002] EWHC 2733 (TCC) at [165], [2003] 1 All ER (Comm) 465 at [166c].

Transport industry

6.137 Software can be manipulated to give whatever reading the writer wishes. Because software is presumed to be 'reliable', software that gives deliberate false data is also presumed to be 'reliable'. It is well known that traffic lights are now generally controlled by software code across a network, and the code can be written in such as way as to break the law. The T-Redspeed traffic light system in Italy is reported to have been developed by Stefano Arrighetti, an engineering student from Genoa. The traffic lights were apparently programmed to remain on amber before turning to red in less than the time set out in regulations.[1]

1 Peter Popham, 'Smart traffic lights rigged to trap drivers', *The Independent* (London, 30 January 2009).

6.138 The 'sudden unintended acceleration' incidents involving the unintended, unexpected and uncontrolled acceleration of modern vehicles with electronic controls raises the issue of the reliability of complex electronic vehicle systems. Consider the prosecution of Ann Diggles, aged 82, who was found not guilty at Preston Crown Court (*R v Ann Diggles* T20157203 before Mr Justice Fraser) for causing death by dangerous driving and death by careless driving when she attempted to park her Nissan Qashqai when it hit and killed Julie Dean, aged 53.[1] The prosecution's case was that the driving of Mrs Diggles caused the accident. The prosecution relied on the evidence from the motor car manufacturer, as reported by the BBC:[2]

> Takuma Nakamura, who is responsible for engine control systems development at Nissan, was asked by prosecutor Richard Archer: 'Is it possible, in your opinion, for a malfunction in an electronic throttle to cause sudden acceleration of the vehicle?'.
>
> Mr Nakamura replied: 'I think that's impossible.'[3]

The expert witness for the defence was Dr Antony F. Anderson CEng FIEE. Dr Anderson pointed out the following:

> A mechanical inspection of the vehicle was carried out. A Nissan garage, on the instruction of the police, downloaded diagnostic trouble codes. The police constable who witnessed the diagnostic testing took a screen shot with his camera that showed three trouble codes. Two of these were past codes of no significance, but one was a current U1000 trouble code. The U1000 code, as I understand it, signifies that there had been a CAN Bus malfunction lasting more than 2 seconds sometime in the ignition cycle during which the incident occurred. Mr Nakamura the senior engineering manager from Nissan Japan, who was sent over to give evidence in the trial, implied that the trouble code was of no significance.[4]

In addition to the evidence from Dr Anderson, two other women came forward at a late stage in the trial to give evidence that they had also had identical experiences. The evidence was that Mrs Diggles and the other two witnesses had their vehicles fully serviced in line with the manufacturer's recommendations.[5] The evidence put before the members of the jury is not readily available, which means it can only be observed that deaths and injuries appear to occur as a result of software failure.[6] It will be of interest to know how the police and prosecution assessed the evidence, including the complexities between the software code and the mechanical and electronic systems. Another prominent example involves Toyota and Lexus motor vehicles, some of which have involved deaths of drivers and their passengers. Michael Barr, in giving expert evidence for the plaintiffs in the case of *Bookout v Toyota Motor Corporation* Case No. CJ-2008-79696[7] stated that:

> A. The Toyota's design actually they have an abysmal design, not just unreasonable in my view, but I use the word abysmal. This was actually the first chapter of my report I wrote because I couldn't believe what I was seeing.
>
> Toyota has a watchdog supervisor design that is incapable of ever detecting the death of a major task. That's its whole job. It doesn't do it. It's not designed to do it.
>
> It also, the thing it does in Toyota's design is lookout for CPU overload, and it doesn't even do that right. CPU overload is when there's too much work in a burst, a period of time to do all the tasks. If that happens for too long, the car can

> become dangerous because tasks not getting to use the CPU is like temporarily tasks dying.
>
> And in Toyota's watchdog you can have any overload going up to one and a half seconds, which at 60 miles an hour I calculated is about the length of a football field, you have any vehicle malfunction for up to a football field in length that's explained only because this watchdog design it [sic] bad, and because the processor is overloaded momentarily. And that should have been also a job of that watchdog supervisor. And that is one they tried to implement and they don't do it well.
>
> They also made a classic blunder, one that's taught by professor like at Dr. Koopman[8] to first year students in his imbedded systems class, which is, you don't dedicate a hardware timer on the main CPU to periodically kick the hardware on the watchdog, because that will keep functioning even though vast portions of the software and the tasks are not rubbing because these interrupts are a higher priority than the tasks.
>
> And so, that is a design that you – and I have spoken about that at many conferences, not doing it that way. And they do that.[9]

1 'Driver cleared over fatal Nissan Qashqai crash', *BBC News* (7 February 2017) <www.bbc.co.uk/news/uk-england-lancashire-38897681>; 'Nissan cars 'sped' without accelerator use, court hears', *BBC News* (6 February 2017), <www.bbc.co.uk/news/uk-england-lancashire-38885809>; 'Driver who killed woman denies mistaking accelerator for brake', *BBC News* (2 February 2017) <www.bbc.co.uk/news/uk-england-lancashire-38846896>.

2 We only have reports from the media to rely on.

3 'Nissan boss denies malfunction caused fatal crash', *BBC News* (31 January 2017), <www.bbc.co.uk/news/uk-england-lancashire-38814890>.

4 Email communication with the author.

5 Gabriella Swerling, ''Runaway car' driver cleared over road death', *The Times* (8 February 2017) 8; James Tozer, 'Great-grandmother who claimed her Nissan Qashqai 'took off' and sped forwards out of control is CLEARED of killing pedestrian', *Daily Mail* (8 February 2017) <www.dailymail.co.uk/news/article-4202212/Woman-claimed-Nissan-Qashqai-took-cleared.html>.

6 David Hechler refers to Betsy Benjaminson, a translator that illustrated the mismatch in evidence when she informed the US authorities: 'Lost in Translation?' *Corporate Counsel* 72 (April 2013) <www.asbpe.org/blog/2014/07/28/david-hechler-wins-asbpes-2014-stephen-barr-award-for-article-on-toyotas-fatal-acceleration-problems/>; see also <http://betsybenjaminson.blogspot.co.uk> for a list of similar incidents across the world; the Crown Price was having troubles with his vehicle, which the manufacturer took pains to resolve: David McNeil, 'Imperial Family's car woes sparked Toyota whistleblower', *The Japan Times* (9 June 2013) <www.japantimes.co.jp/news/2013/06/09/business/corporate-business/imperial-familys-car-woes-sparked-toyota-whistleblower/#.WJ14B-l4j8s>.

7 The trial was held in the District Court of Oklahoma County State of Oklahoma before the Hon Patricia G. Parrish, District Judge; Transcript (not proofread) of the trial 14 October 2013 (Reported by Karen Twyford, RPR): examination and cross examination of Michael Barr, available at <www.safetyresearch.net/Library/Bookout_v_Toyota_Barr_REDACTED.pdf>.

8 Dr Koopman is an Associate Professor at Carnegie Mellon University, Department of Electrical and Computer Engineering.

9 At 70–1. Professor Philip Koopman also gave evidence in this case, and his assessment of the problem was similar to that of Mr Barr, for which see <www.safetyresearch.net/blog/articles/toyota-unintended-acceleration-and-big-bowl-"spaghetti"-code>.

6.139 Software in vehicles can also be manipulated to give the false assurance of regulatory compliance. In September 2015, the United States Environmental Protection Agency issued a notice of violation of the Clean Air Act to Volkswagen AG, Audi AG, and Volkswagen Group of America, Inc.[1] The notice alleged that four-cylinder Volkswagen and Audi diesel cars covering the years 2009–15 include software that circumvented the emissions standards for some air pollutants. The state of California

Air Resources Board had issued a separate In-Use Compliance letter to Volkswagen,[2] and the two agencies initiated investigations based on the allegations. A software algorithm on certain Volkswagen vehicles switched the full emissions controls on only when the car detected as undergoing official emissions testing. Thus the effectiveness of the emission control devices was greatly reduced during normal driving. This meant that motor vehicles met the emissions standards in the laboratory or testing station, but during normal operation, the vehicles emitted nitrogen oxides, or NOx, at up to 40 times the standard. Over a one year period of operation, the emission of this extra pollutant by Volkswagen was estimated to have resulted in 5 to 50 premature deaths.[3] The Department of Justice subsequently filed a complaint for alleged violations of the Clean Air Act.[4]

1 For details, see <www.epa.gov/vw/learn-about-volkswagen-violations>.

2 Letter from the Air Resources Board to Volkswagen AG, Audi AG, and Volkswagen Group of America, Inc dated 18 September 2015, reference number IUC.2015-007, available at <www.arb.ca.gov/newsrel/in_use_compliance_letter.htm>.

3 Lifang Hou, Kai Zhang, Moira A. Luthin and Andrea A Baccarelli, 'Public health impact and economic costs of Volkswagen's lack of compliance with the United States' emission standards' (2016) 13 International Journal of Environmental Research and Public Health 891; Gregory J Thompson, Daniel K Carder, Marc C Besch, Arvind Thiruvengadam and Hemanth K Kappanna, *Final Report: In-Use Emissions Testing of Light-Duty Diesel Vehicles in the United States* (Center for Alternative Fuels, Engines & Emissions Department of Mechanical & Aerospace Engineering, West Virginia University, 15 May 2014) <www.eenews.net/assets/2015/09/21/document_cw_02.pdf>.

4 Press release: 'United States Files Complaint Against Volkswagen, Audi and Porsche for Alleged Clean Air Act Violations', Monday, January 4, 2016, at <www.justice.gov/opa/pr/united-states-files-complaint-against-volkswagen-audi-and-porsche-alleged-clean-air-act>, including a link to the original Complaint; an amended Complaint was submitted on 7 June 2016 and is available at <www.epa.gov/sites/production/files/2016-10/documents/amendedvw-cp.pdf>.

Emergency services: London Ambulance computer aided dispatch system

6.140 In 1992, the London Ambulance computer aided dispatch system failed. A complex set of circumstances resulted in an effective failure of the dispatching system, which are set out in paragraph 1996 of the Report.[1] Apparently 'the computer system itself did not fail in a technical sense. … However, much of the design had fatal flaws that would, and did, cumulatively lead to all of the symptoms of systems failure'.[2] Among the contributing factors were 'exception messages' and 'requests for attention' which scrolled off the screen because of the large number of messages generated.[3] There is also a suggestion that one member of staff was not using the system as expected,[4] and the problems were compounded by 'a genuine failure of crews to press the correct status button owing to the nature and pressure of certain incidents'.[5] This was so even though the individuals that used the new system were from a skilled and trained pool of staff, namely ambulance crews and controllers. Other problems have occurred since.[6]

1 *Report of the Inquiry into the London Ambulance Service* (South West Thames Regional Health Authority, 1993) – a scanned version is available at <http://www0.cs.ucl.ac.uk/staff/A.Finkelstein/las.html>; P Mellor, 'CAD: Computer-Aided Disaster' (1994) 1 High Integrity Systems Journal 101; Anthony Finkelstein and John Dowell, 'A comedy of errors: the London Ambulance Service case study', *Proceedings of the 8th International Workshop on Software Specification & Design IWSSD-8* (IEEE CS Press 1996), 2–4; Paul Beynon-Davies, 'Information systems "failure" and risk assessment: the case of the London ambulance service computer and despatch system', in G Doukidid, B Galliers, H Kremar and F. Land (eds), *Proceedings of the 3rd European Conference on Information Systems* (Athens, 1–3 June

1995), pp. 1153– 70; Paul Beynon-Davies, 'Human error and information systems failure: the case of the London ambulance service computer-aided despatch system project' (1999) 11 Interacting with Computers 699; D Dalcher, 'Disaster in London: The LAS Case study' [1999] Engineering of Computer-Based Systems 41.

2 *Report of the Inquiry into the London Ambulance Service*, para 1007(x).

3 *Report of the Inquiry into the London Ambulance Service*, paras 4012(c) and 4023.

4 *Report of the Inquiry into the London Ambulance Service*, para 4025.

5 *Report of the Inquiry into the London Ambulance Service*, para 4009(b).

6 Kelly Fiveash, 'London Ambulance Service downed by upgrade cockup', *The Register* (9 June 2011); Jon Ironmonger, 'Ambulance system failure 'might have led to patient death', *BBC News* (6 January 2017).

Medical industry

6.141 The widespread use of computer devices in the medical industry has also given rise to incidents where the reliability of devices and software has been called into question. For instance, patients have been affected by an error in clinical IT software,[1] and one study of a hospital computerized physician order entry systems in the United States of America illustrated a number of errors that the system was supposed to resolve, such as an increased probability of prescribing errors. There were 12 flaws in the interface used by humans that reflected machine rules that in turn did not correspond to how work was organized or the usual behaviour of those using the system.[2] There is an increasing volume of articles on this topic,[3] and it would appear that some, and not all, of the problems were due to software defects,[4] but it is now very clear that software helps kill people in hospitals.[5]

1 Alex Matthews-King, 'GPs told to review patients at risk as IT error miscalculates CV score in thousands', *Pulse Today* (11 May 2016) <www.pulsetoday.co.uk/your-practice/practice-topics/it/gps-told-to-review-patients-at-risk-as-it-error-miscalculates-cv-score-in-thousands/20031807.fullarticle>.

2 Ross Koppel, Joshua P Metlay, Abigail Cohen, Brian Abaluck, A Russell Localio, Stephen E Kimmel and Brian L Strom, 'Role of computerized physician order entry systems in facilitating medication errors' (2005) 293 Journal of the American Medical Association 1197.

3 E Alberdi, A A Povyakalo, L Strigini and P Ayton, 'Computer aided detection: risks and benefits for radiologists' decisions', in E Samei and E Krupinski (eds), *The Handbook of Medical Image Perception and Techniques* (Cambridge University Press 2009), 320–32.

4 Frances E Zollers, Andrew McMullin, Sandra N Hurd and Peter Shears, 'No more soft landings for software: liability for defects in an industry that has come of age' (2004) 21 Santa Clara High Tech LJ 745; Sharona Hoffman and Andy Podgurski, 'E-Health hazards: provider liability and electronic health record systems (2009) 24 Berkeley Tech LJ, 1523; Paul T Lee, Frankie Thompson and Harold Thimbleby, 'Analysis of infusion pump error logs and their significance for health care' (2012) 21 British Journal of Nursing S12; John M Curran and Mark A Berman, 'Gremlins and glitches: using electronic health records at trial' (2013) 85 New York State Bar Journal 20; Courtney L Davenport, 'Dangers of electronic medical systems' (2013) 49 Trial: The National Legal Newsmagazine 14; Timothy P. Blanchard and Margaret M Manning, 'Electronic medical record documentation: inherent risks and inordinate hazards', in Alice G Gosfield (ed.), *Health Law Handbook* (Thompson Reuters 2016), pp. 246–97; *Karam v. Adirondack Neurosurgical Specialists, P.C.*, 93 A.D.3d 1260 (2012), 941 N.Y.S.2d 402, 2012 N.Y. Slip Op. 02182 (evidence pointed to error in software), motion for reargument or leave to appeal to the Court of Appeals denied, 96 A.D.3d 1513 (2012), 945 N.Y.S.2d 588, 2012 N.Y. Slip Op. 04645, motion for leave to appeal denied, 19 N.Y.3d 812 (2012), 976 N.E.2d 251, 951 N.Y.S.2d 722, 2012 N.Y. Slip Op. 83806.

5 Yong Y Han, Joseph A Carcillo, Shekhar T Venkataraman, Robert S B Clark, R Scott Watson, Trung C. Nguyen, Hülya Bayir and Richard A Orr, 'Unexpected increased mortality after implementation of a commercially sold computerized physician order entry system' (2005) 116 Pediatrics 1506; Harold Thimbleby, 'Ignorance of interaction programming is killing people', *Interactions* (September and October 2008), 52.

Banking industry

6.142 The presumption that computers are reliable is particularly relevant with regard to banking. Banks across the world have introduced very complex systems and networks to control the flow of transactions, many of which are no longer under the sole control of the banks themselves. That a bank benefits from the presumption that its computers and networks, including the computers and networks it relies upon over which it has no direct control were in order at the material time, puts an impossible burden on the customer. For a customer in dispute with his bank to challenge this presumption, he will require significant knowledge of the computers, systems and networks operated by the bank, how they work, and where the vulnerabilities might lie, including the results of relevant audits, both internal and external – a task well beyond the vast majority of customers, including most lawyers without the benefit of expert advice, which in itself is difficult to obtain.

6.143 This is a problem that arose in the case of *Post Office Ltd v Castleton*.[1] The Post Office took action to recover monies on an account stated by one of its former subpostmasters, Mr Castleton. Havery J had the task of making a judgment in the absence of legal representation on behalf of the defendant. Mr Castleton challenged the reliability of the Horizon computer system, a computer system that is centrally controlled. There was no evidence before the court relating to the software code relied upon by the Post Office, and one knowledgeable witness, Anne Chambers, a system specialist employed by Fujitsu, gave evidence of the system. In cross examination, Mrs Chambers admitted the system caused losses at a branch at Callender Square in Falkirk, although there was no record as to the basis of her knowledge, but asserted that such a problem did not occur at the branch at which Mr Castleton worked. The judge had to reach a decision with the evidence before him. Some of the evidence appeared to support Mr Castleton's assertion that the monies lost were caused by software errors: unexplained entries like 'Declare stamp total £0.00' and 'Declare cash £0.00', an apparent error in the figure of £3,533.30, and errors in intermediate figures such as £2,654.60. The judge noted these errors, but assumed these had no apparent effect on the accounts in question, and did not consider there was 'exiguous evidence that the Horizon system was flawed'.[2] Although the judge was satisfied that an intermediate figure of £2,654.60 could not be right, without any evidence on the point, the judge offered the explanation that it 'may be a mistyping of the entry into the computer', which in his view did not affect the weekly accounts, and he noted that this error had not been put forward as evidence of a fault in the Horizon system.[3] Notwithstanding these errors, the judge decided that there were substantial real unexplained deficiencies that provided irrefutable evidence that the post office under Mr Castleton's supervision was not properly managed at the material time, and that the losses must have been caused by his own error, or that of his assistants. The Post Office was therefore entitled to summarily terminate his contract.[4]

1 [2007] EWHC 5 (QB). The Post Office Horizon system has been the topic of particular scrutiny, for which see the editorial in the 2015 issue of the Digital Evidence and Electronic Signature Law Review and the introductory remarks to the transcript of the trial of *R v Seema Misra* (October 2010), reproduced in full in the same journal. A number of prosecutions by the Post Office regarding the Horizon system are the subject of a review by the Criminal Cases Review Commission, and there is a class action in the civil courts against the Post Office (*Bates and others v Post Office Limited*, Case no HQ16X01238). The first hearing was heard before Senior Master Fontaine in the High Court of Justice, Queen's Bench Division, on 26 January 2017.

2 [2007] EWHC 5 (QB), [15]. In *Banks v Revenue & Customs* [2014] UKFTT 465 (TC), in response to the appellant's assertions that the online process for submitting tax forms was flawed, Revenue and Customs rejected the claim without providing any evidence, the members of the tribunal reporting, at [22], that 'HMRC says that it interrogated its computer system, and found no faults'. In addition, the members of the tribunal stated at [28], in the absence of any evidence to make such an assessment, that 'It is equally difficult to envisage HMRC's systems failing in such a rudimentary way'.
3 [2007] EWHC 5 (QB), [20].
4 [2007] EWHC 5 (QB), [40].

6.144 Issues regarding the reliability of banking systems manifested themselves in the problems in the UK in June and July 2012 with RBS, NatWest and Ulster banks.[1] On 19 June 2012, an important item of software known as CA-7 was updated. This software controls the batch processing systems that deal with retail banking transactions. It is used to automate large sequences of batch mainframe work, usually referred to as 'jobs'. The jobs take transactions from various places, such as ATM withdrawals, automatic salary payments and such like, so that accounts are credited and debited with the correct amounts by the next morning. The software initiates jobs, and when one job is finished, a new job will be initiated. Accounts are processed overnight when the mainframes are less busy, and finish by updating the master copy of the account in a system known as Caustic. It appears that the update made to CA-7 caused the files to run incorrectly or not to run at all for three nights. David Silverstone, delivery and solutions manager for NMQA, which provides automated testing software to a number of banks, is quoted to the effect that such problems can always be avoided if there is sufficient testing of the update before it is put into operational use.[2] Michael Allen, director of IT service management at Compuware, is reported to have said:

> The problem is that IT systems have become vastly more complex. Delivering an e-banking service could be reliant on 20 different IT systems. If even a small change is made to one of these systems, it can cause major problems for the whole banking service, which could be what's happened at NatWest. Finding the root cause of the problem is probably something NatWest is struggling with because of the complexity of the IT systems in any bank.[3]

1 For detailed information, the reader is directed to the Treasury Select Committee web page on the Parliament website.
2 Charles Arthur, 'How NatWest's IT meltdown developed', *The Guardian* (25 June 2012).
3 Anna Leach, 'Natwest, RBS: when will bank glitch be fixed? Probably not today', *The Register* (22 June 2012).

6.145 The complexity of the problem is highlighted in an article written by Hilary Osborne in *The Guardian* in 2014,[1] in which the issues were explained:

> 'The banks do have a problem, but it's not a new problem, and it's not an easy problem to fix, which is why it's taking so long,' says David Bannister, editor of Banking Technology magazine. 'In the old days these machines just had to run overnight in batch mode – it was like newspapers with just one edition – but now they have to deal with news that is being updated throughout the day. The users – us – are using internet banking, ATMs, we're spending money online. The reconciliation between what is going on in the background is the hard part, and the gulf is widening all the time.'
>
> Ben Wilson, associate director of financial services for techUK, says some of the 'legacy systems' at banks are 30–40 years old and were originally set up for branch banking, but 'then they needed to be ATM-focussed, then there was online

> banking, then mobile banking'. He says: 'Banks have bolted on these changes because it is cheaper and less risky than starting from scratch, but every time you bolt on a change it becomes more complex.'
>
> As well as new banking channels, systems are also tinkered with whenever regulatory changes are made, and when a product is withdrawn or changed.
>
> Jim McCall, managing director of the Unit, which works with banks and other companies on their mobile apps, says that while anyone now building a system from scratch would 'abstract out as much as possible so [different elements] are not as reliant on each other', the banks' systems often resemble a house of cards. 'If you make a change to a tiny bit of code on one thing it is like the butterfly flapping its wings far away and somewhere someone's mobile app stops working,' he says.
>
> To make things more complicated, says Colin Privett, UK managing director of software firm Cast, new functions are usually 'written in different programming languages, on different machines, by different teams'. He adds: 'This prevents a single person/team from ever fully understanding the entire structure of a system. That is why when things do go wrong it can often take hours, or even days, to fix as teams scramble to find out where the problem lies.'

1 Hilary Osborne, 'Why do bank IT systems keep failing?', *The Guardian* (27 January 2014).

6.146 The effects of the CA-7 imbroglio were considerable. In some cases people were left homeless after the computer problems meant house purchases fell through; others were stranded abroad, unable to obtain access to funds which should have been in their account; wages and direct debits were not paid, and it is reported that one person spent the weekend in prison because the computer failure meant his bail money was not processed.[1] The problems continued into 2014.[2] In December 2014, the Royal Bank of Scotland Plc, National Westminster Bank Plc and Ulster Bank Ltd faced a combined financial penalty of £42m by the Financial Conduct Authority for breaches of Principle 3 of the 'principles for businesses', forming part of 'The principles of good regulation', which requires a firm to take reasonable care to organize and control its affairs responsibly and effectively with adequate risk management systems,[3] and the Prudential Regulation Authority imposed a financial penalty of £14m on the same banks for their failure to meet their obligations to have adequate systems and controls to identify and manage their exposure to IT risks.[4]

1 James Hall and Gordon Rayner, 'RBS computer failure condemns man to spend weekend in the cells', *The Telegraph* (25 June 2012).
2 Emma Dunkley, 'RBS and NatWest to plough £1bn into digital upgrade', *Financial Times* (28/29 June 2014), p. 18.
3 <www.fca.org.uk/static/documents/final-notices/rbs-natwest-ulster-final-notice.pdf>.
4 <www.bankofengland.co.uk/pra/Documents/supervision/enforcementnotices/en201114.pdf>.

6.147 The problem of complexity and the difficulties in understanding and maintaining another banking system were emphasized in the report by Deloitte of the failure of the Real-Time Gross Settlement (RTGS) system operated by the Bank of England in 2014.[1] The report stated:

> 133. During the 18 years since RTGS was first launched, the incremental changes have resulted in an increase in complexity and a system which is now more difficult to understand and maintain. In particular, the LSM and MIRS changes introduced additional functionality with an associated increase in complexity.
>
> 134. In combination with the ageing development language used to program

> RTGS, the result is a system which is more complex to support, heavily reliant on the skills and experience of the team to support it, and more susceptible to errors which take longer to diagnose. Therefore there is an increased risk of functional or configuration changes causing errors and if or when the system does fail it may take longer to resolve the issue.

1 Deloitte, *Independent Review of RTGS Outage on 20 October 2014* (23 March 2015), available at <www.bankofengland.co.uk/publications/Documents/news/2015/rtgsdeloitte.pdf>.

6.148 In this case, there was a design defect. The defect was mentioned at paragraph 151 of the report, but it had been redacted to such an extent that there was no meaningful text. The only information available is that a process known as 'Process A functionality' was changed in April 2014 and tested in May 2014 in preparation for the anticipated transfer of CHAPS members, and a design defect was introduced at this stage. This was the cause of the failure.[1]

1 Independent review of RTGS outage on 20 October 2014: Bank of England's response, at <www.bankofengland.co.uk/publications/Documents/news/2015/rtgsresponse.pdf>.

6.149 Other examples include Deutsche Bank AG, where a coding error caused Deutsche to reverse the buy/sell indicator for its CFD Equity Swaps in 2013. This meant it reported them inaccurately to the Financial Conduct Authority (FCA). The FCA imposed a financial penalty of £4,7818,800 on Deutsche for failing to provide accurate reports in accordance with the provisions of the Markets in Financial Instruments Directive.[1] In 2014, the Co-operative Bank identified that statements on a number of loans had been issued three days late because of a software error. Under the provisions of s 6 of the Consumer Credit Act 2006, which inserted s 77A into the Consumer Credit Act 1974, it is necessary to provide an annual statement to each borrower for a fixed-sum credit agreement, which should set out the amount borrowed, the money paid, the interest and the outstanding amount. If the creditor fails to provide the debtor with an annual statement, the creditor is not entitled to enforce the agreement during the period of the failure to comply, and the debtor is not liable to pay any interest during the period. The bank set aside £109.5m to refund interest payments for this breach of the Act.[2]

1 <https://www.fca.org.uk/static/documents/final-notices/deutsche-bank-ag-2015.pdf>; Directive 2004/39/EC of the European Parliament and of the Council of 21 April 2004 on markets in financial instruments amending Council Directives 85/611/EEC and 93/6/EEC and Directive 2000/12/EC of the European Parliament and of the Council and repealing Council Directive 93/22/EEC, (OJ L 145, 30.4.2004, p.1).
2 Adam Leyland and Beth Brooks, 'The Co-operative Bank's £400m costs bill caused by 'programming error'', *The Grocer* (29 March 2014), at <www.thegrocer.co.uk/channels/supermarkets/the-co-operative-group/programming-error-to-blame-for-co-op-banks-400m-bill/356022.article>; The Co-operative Bank plc, Annual report and accounts for 2013, 151 section 2(iv).

Interception of communications

6.150 In the half-yearly report in July 2015, the Report of the Interception of Communications Commissioner illustrated the effect that errors in software code had in the interception of communications.[1] Although the number of technical errors were low in comparison to the overall number of requests made, nevertheless the effect such errors has on innocent parties is significant. In paragraph 5.28, it was indicated that eight of ten errors made in relation to resolving IP addresses to individuals related to

investigations into the sexual exploitation of children or cases where serious concerns were raised in relation to the welfare of a child.[2] The Commissioner commented, at paragraphs 5.29 and 5.37:

> Regrettably when errors occur in relation to the resolution of IP addresses the consequences are particularly acute. An IP address is often the only line of enquiry in a child protection case (so called 'single strand' intelligence), and it may be difficult for the police to corroborate the information further before taking action. Any police action taken erroneously in such cases, such as the search of an individual's house who is unconnected with the investigation or a delayed welfare check on an individual whose life is believed to be at risk, can have a devastating impact on the individuals concerned.
>
> …
>
> 5.37 … The eight technical system errors led to four warrants being executed at premises unconnected with the investigations and in one of these instances an individual was arrested. In another case the error delayed a welfare check on a child believed to be in crisis. In one instance a person unconnected with the investigation was visited by police. The majority of these errors resulted in communications data being obtained in relation to individuals who were unconnected with those investigations.

1 The Rt Hon. Sir Anthony May, *Half-yearly report of the Interception of Communications Commissioner* (July 2015, HC 308, SG/2015/105).
2 There is no suggestion from these examples that it was in error. The report may mean errors in resolving IP addresses in criminal investigations.

6.151 In his Report, the Commissioner said that the Crown Prosecution Service used funds provided by the government to work with vendors and the Home Office to develop secure disclosure systems – and although money has been spent on this issue, nevertheless, technical issues continue to arise.[1] As a result of the disclosure of the technical errors, the Commissioner made a number of recommendations regarding technical system errors:

> 11 Ensure that the [Communication Service Provider] CSP secure disclosure systems are tested sufficiently prior to implementation and after significant updates or upgrades.
>
> 12 Ensure there is standardization and as much consistency as possible in relation to the data entry requirements on the different CSP secure disclosure systems.
>
> 13 Requirement for [Single Point of Contact] SPoC to inform CSP immediately if an error is identified which might be the result of a technical system fault (even where the error has been classified as a recordable error).
>
> 14 Ensure that there are regular quality assurance audits of the CSP secure disclosure systems to identify any faults at an earlier stage.
>
> 15 Ensure that the CSPs and system vendors are aware of the potential significant consequences of system errors, that the public authorities are informed of any systems errors immediately and the errors are fixed at the earliest opportunity.[2]

1 At paragraph 5.53.
2 At paragraph 5.40.

Most computer errors are either immediately detectable or result from input errors

6.152 Let us consider the proposition that most computer errors are either immediately detectable or result from errors in the data entered into the machines. The evidence is to the contrary: Mr Adams demonstrated that a third of software faults in a large IBM study took at least 5,000 execution years to appear for the first time (this was one of the largest studies of all time);[1] Professor Les Hatton and Andy Roberts conducted a study that demonstrated that seismic programs developed by oil companies were shown to have been used for many years even though they were defective;[2] and Nancy G. Leveson and Clark S. Turner demonstrate that between June 1985 and January 1987, the Therac-25 medical linear accelerator was involved in massive radiation overdoses, causing the deaths of six people, while others were seriously injured. The detailed investigations eventually indicated that the main cause of the deaths was software errors. Some of the lessons gleaned from the work by Nancy Leveson included: too much confidence was placed in the software, an assumption by lay people that software will not or cannot fail, and engineers ignoring software when analysing faults, because it was assumed the hardware was at fault, not the software.[3] In this respect, opinions have not changed since 1987.[4] Toyota, when investigating sudden unintended acceleration in some of its motor cars in the US, did not include software engineers in its investigations, and incorrectly ruled out software as the cause of the deaths and injury of people.[5]

1 Edward N Adams, 'Optimizing preventive service of software products' (1984) 28 IBM Journal of Research and Development 2.

2 Les Hatton and Andy Roberts, 'How accurate is scientific software?' (1994) 20 IEEE Transactions on Software Engineering 785.

3 'An investigation of the Therac-25 accidents' (1993) 26 Computer 18 (note the additional information in Nancy Leveson, *Software, System Safety and Computers* (Addison-Wesley 1995); for descriptions of what some of the patients suffered, see Leonard Lee, *The Day The Phones Stopped: The Computer Crisis – The What and Why of It, and How We Can Beat It* (Donald I. Fine,1991), ch. 1.

4 Simon Oxenham, 'Thousands of fMRI brain studies in doubt due to software flaws', *New Scientist* (18 July 2016), at <www.newscientist.com/article/2097734-thousands-of-fmri-brain-studies-in-doubt-due-to-software-flaws/>; Anders Eklund, Hans Knutsson and Thomas Nichols, 'Cluster failure: why fMRI inferences for spatial extent have inflated false-positive rates' (2016) 113 Proceedings of the National Academy of Sciences of the United States of America 7900–5 <www.pnas.org/content/113/28/7900.full.pdf>.

5 Transcript (not proofread) of *Bookout v Toyota Motor Corporation Case* No. CJ-2008-7969 (Reported by Karen Twyford, RPR): examination and cross examination of Michael Barr 14 October 2013, 76–7, available at <www.safetyresearch.net/Library/Bookout_v_Toyota_Barr_REDACTED.pdf>.

6.153 Uncovering the faults in devices controlled by software used in medicine is now considered to be an important research area,[1] and in November 2000, 28 patients at the National Cancer Institute in Panama were given massive overdoses of gamma rays partly due to limitations of the computer program that guided use of a radiation-therapy machine. A number of patients died.[2]

1 Kevin Fu, 'Trustworthy medical device software' (comprising Appendix D, 97–118) in Thereza Wizemann (ed.), *Public Health Effectiveness of the FDA 510(k) Clearance Process: Measuring Postmarket Performance and Other Select Topics: Workshop Report* (Food and Drug Administration 2011), available at <www.ncbi.nlm.nih.gov/books/NBK209656/>; see also Senate Hearing 112-92, United States Senate, Hearing on a Delicate Balance: FDA and the Reform of the Medical Device Approval Process,

13 April 2011, at <www.aging.senate.gov/hearings/a-delicate-balance-fda-and-the-reform-of-the-medical-device-approval-process>.

2 Deborah Gage and John McCormick, *We Did Nothing Wrong: Case 109 A Dissection*, at <http://disciplinas.stoa.usp.br/pluginfile.php/31797/mod_resource/content/1/casoCancerPanama.pdf>; *Investigation of an Accidental Exposure of Radiotherapy Patients in Panama Report of a Team of Experts* (International Atomic Energy Agency, 26 May–1 June 2001), at <www-pub.iaea.org/mtcd/publications/pdf/pub1114_scr.pdf>; Cari Borrás, 'Overexposure of radiation therapy patients in Panama: problem recognition and follow-up measures' (2006) 20 Revista Panamericana de Salud Pública/Pan American Journal of Public Health 173.

6.154 The observations by Professor Leveson will invariably remain relevant: the Toyota recall exercise in late 2009 and early 2010 serves to illustrate this point.[1] The US Congressional Committee on Energy and Commerce heard evidence on this matter, and a report by The National Highway Traffic Safety Administration and the National Aeronautics and Space Administration (NHTSA-NASA), which conducted a study into the problem entitled 'Study of unintended acceleration in Toyota vehicles', a revised version of which was published on 15 April 2011,[2] concluded that it was not proven that faulty software caused the problems, although it was accepted that just because no software faults could be found did not mean that software faults did not occur. The methods used to investigate this matter were challenged.[3]

1 A number of motor manufacturers are facing similar legal actions. It was known that sudden acceleration occurred in the 1980s and 1990s, for which see James Castelli, Carl Nash, Clarence Ditlow and Michael Pecht, *Sudden Acceleration: The Myth of the Driver Error* (Calce EPSC Press 2003).

2 Available at <www.nasa.gov/topics/nasalife/features/nesc-toyota-study.html>.

3 For which see Michael Barr, 'Firmware forensics: best practices in embedded software source code discovery' (2011) 8 Digital Evidence and Electronic Signature Law Review 148. For an earlier article, see Joel Finch, 'Toyota sudden acceleration: a case study of the National Highway Traffic Safety Administration recalls for change' (2010) 22 Loy Consumer L Rev 472.

6.155 Civil proceedings were subsequently initiated by a number of people across the US. In *Bookout v Toyota Motor Corporation* Case No. CJ-2008-7969,[1] Michael Barr, an expert in embedded software, gave evidence for the plaintiff regarding the software code in the relevant motor vehicles. He was also cross examined about aspects of the NHTSA Report among other issues. His evidence demonstrated that there were a significant number of errors in the software (referred to as 'bugs' in the transcript):

> Q. Did you find all the bugs in the software that you reviewed?
>
> A. Absolutely not.
>
> Q. Why not?
>
> A. Because there is a lot of bugs, and all indications are that there are many more. We haven't specifically gone out looking for bugs. The metrics, like the code complexity and a number of global variables, indicate the presence of large numbers of bugs. And just the overall style of the coded is suggestive that there will be numerous more bugs that we haven't found yet.[2]

1 The trial was held in the District Court of Oklahoma County State of Oklahoma before the Hon Patricia G. Parrish, District Judge.

2 Transcript (not proofread) of the trial 14 October 2013 before the Hon Patricia G. Parrish, District Judge (Reported by Karen Twyford, RPR): examination and cross examination of Michael Barr, 47–8, available at <www.safetyresearch.net/Library/Bookout_v_Toyota_Barr_REDACTED.pdf>.

6.156 He also demonstrated that motor cars are largely run by software now. In fact, motor cars have more software code than aircraft, and are prone to software recalls.[1]

Drivers no longer have total control over their vehicles. For instance, it was explained how the driver is no longer in direct control of the throttle:

> … But the driver had always been directly in control of the air, which is directly related to how much power the engine has. When electronic throttle control comes in, you have software that is now responsible for all three of them at once. So you have a portion of the software, the job of which is to make the spark at the right time, inject the fuel at the right time and the right amount, and open the throttle a certain amount.
>
> …
>
> The software in electronic throttle control is responsible for all three things, which means if the software malfunctions, it has control of the engine and can take you for a ride. What is of particular importance is that there is another part of the software that is looking at the driver controls, looking at the accelerator pedal and cruise control -- it is looking at more than that, but that is a simplification, that is appropriate right now -- so there is a part of the software looking at what the accelerator pedal position is, is it down, is it up, how much down. Then that is translating that into a calculated throttle angle. And then another part of the software is performing the sparking and the throttle control.[2]

1 R.N. Charette, 'This car runs on code', *IEEE Spectrum* (1 February 2009) <http://spectrum.ieee.org/transportation/systems/this-car-runs-on-code>; Mössinger, 'Software in automotive systems' 92.
2 Transcript of the trial 14 October 2013, 53.

6.157 Mr Barr established that the motor vehicle had errors in the throttle system:

> A. So the first main conclusion is that the 2005 Camry electronic throttle control, the software is of unreasonable quality. It contains bugs, but that's not the only reason it is of unreasonable quality. And it's otherwise defective for a number of reasons. This includes bugs that when put together with the defects can cause unintended acceleration.
>
> Q. As we go forward are you going to explain to us how those problems that you found will cause an unintended acceleration?
>
> A. Yes.
>
> Q. Then you mentioned the code quality metrics. What do you mean about that?
>
> A. So the code complexity and the McCabe Code Complexity is one of the measures of that.[1] And the code complexity for Toyota's code is very high. There are a large number of functions that are overly complex. By the standard industry metrics some of them are untestable, meaning that it is so complicated a recipe that there is no way to develop a reliable test suite or test methodology to test all the possible things that can happen in it. Some of them are even so complex that they are what is called unmaintainable, which means that if you go in to fix a bug or to make a change, you're likely to create a new bug in the process. Just because your car has the latest version of the firmware – that is what we call embedded software – doesn't mean it is safer necessarily than the older one.[2]

1 McCabe Code Complexity has no sound theoretical basis. It is a rule of thumb. I owe this point to Dr Michael Ellims.
2 Transcript of the trial 14 October 2013, 65–6.

6.158 It was Mr Barr's opinion that 'ultimately my conclusion is that this Toyota electronic throttle control system is a cause of UA software malfunction in this electronic throttle module, can cause unintended acceleration'.[1] The members of the jury found in favour of the plaintiffs and awarded damages of US$1.5m to each of the plaintiffs.

The US Department of Justice subsequently concluded a criminal investigation into the Toyota Motor Company regarding the widespread incidents of unintended vehicle acceleration that caused panic for Toyota owners between 2009 and 2010. It was established with certainty that Toyota intentionally concealed information and misled the public about the safety issues behind these recalls.[2] It was alleged that Toyota made misleading public statements to consumers and gave inaccurate facts to Members of Congress and concealed the extent of problems that some consumers encountered from federal regulators. In its settlement with the Department of Justice, Toyota admitted its wrongdoing in making such misleading statements in the Statement of Facts filed with the criminal information, and also admitted that it undertook these actions as an act of concealment as part of efforts to defend its brand. In consequence, Toyota paid a financial penalty of US$1.2 billion under the settlement.[3]

1 Transcript of the trial 14 October 2013, 67.

2 The literature on this topic in general merits further analysis, but is beyond the scope of this chapter: Suzanne M Kirchhoff and David Randall Peterman, 'Unintended Acceleration in Passenger Vehicles' (Congressional Research Service 7-5700, R41205, 26 April 2010); R Graham Esdale Jr and Timothy R Fiedler, 'Toyota's deadly secrets', 46-SEP Trial 16; Finch, 'Toyota sudden acceleration', 472; Molly S O'Neill, 'Faulty cars or faulty drivers: the story of sudden acceleration and Ford Motor Company', (undated and scanned images from an unidentified book), available at <www.suddenacceleration.com/article-2/>; Scott Elder and Travis Thompson, 'Recent development in automobile consumer class actions', 41-FALL Brief 44; Katherine Gardiner, 'Recent developments in automobile law' (2011–12) 47 Tort Trial & Insurance Practice Law Journal 45; Joseph Gavin, 'Crash test dummies: what drives automobile safety in the United States?' (2012) 25 Loy Consumer L Rev 86; Maria N Maccone, 'Litigation concerning sudden unintended acceleration', 132 Am Jur Trials 305; Qi Van Eikema Hommes, 'Review and Assessment of the ISO 26262 Draft Road Vehicle – Functional Safety' (SAE Technical Paper 2012-01-0025, 2012); David C Vladeck, 'Machines without principals: liability rules and artificial intelligence' (2014) 89 Washington Law Review 117; Aaron Ezroj, 'Product liability after unintended acceleration: how automotive litigation has evolved' (2014) 26 Loy Consumer L Rev 470; Anthony F Anderson, 'Intermittent electrical contact resistance as a contributory factor in the loss of automobile speed control functional integrity' (2014) 2 IEEE Access 258; Anthony F Anderson, 'Case study: NHTSA's denial of Dr Raghavan's petition to investigate sudden acceleration in Toyota vehicles fitted with electronic throttles' (2016) 4 IEEE Access 1417.

3 <www.justice.gov/usao-sdny/programs/victim-witness-services/united-states-v-toyota-corporation>.

Challenging the authenticity of digital data – trial within a trial

6.159 Laying the evidentiary foundations for the authenticity of digital evidence is discussed elsewhere in this text, but if the authenticity of evidence is raised by one of the parties, it is appropriate to deal with it in a trial within a trial.[1] This will be a rare occurrence, as noted by Bedlan J in *R v Wayte (William Guy Alexander)*:

> It may be that in very rare cases, there will have to be a trial within a trial on the issue of the admissibility … but on such an issue, where the party producing the document and arguing for its admissibility contends that it is genuine … the issue will invariably be left to the jury …[2]

1 R. Pattenden, 'Pre-verdict judicial fact-finding in criminal trials with juries' (2009) 29 Oxford Journal of Legal Studies 1.

2 (1982) 76 Cr App R 110, CA, 118.

6.160 In *R v Stevenson*,[1] Kilner Brown J was required to establish whether audio tapes

were originals. After a lengthy and careful examination of the evidence held in a trial within a trial, it became clear that there was an opportunity for someone to have interfered with the original tape, and there was evidence that some interference might have taken place. Given the nature of the evidence before him, he said:

> Once the original is impugned and sufficient details as to certain peculiarities in the proffered evidence have been examined in court, and once the situation is reached that it is likely that the proffered evidence is not the original, is not the primary and best evidence, that seems to be to create a situation in which, whether on reasonable doubt or whether on a prima facie basis, the judge is left with no alternative but to reject the evidence.[2]

1 [1971] 1 All ER 678, [1971] 1 WLR 1.
2 [1971] 1 All ER 678, [1971] 1 WLR 1, 3G.

6.161 In the case of *R v Robson (Bernard Jack); R v Harris (Gordon Federick)*[1] the defence raised the issue of the admissibility of the evidence of 13 tape recordings. The judge had to consider whether, on the face of it, the tapes were authentic in the absence of the members of the jury. Shaw J heard evidence in a trial within a trial from a number of witnesses who gave evidence of the history of the tapes, from the actual process of recording to the time they were produced in court. He also listened to four experts called on behalf of the defence, whose examination of the tapes led them to question their originality and authenticity. The prosecution called a separate witness in rebuttal. After hearing the evidence, Shaw J decided that the tape recordings were originals and authentic, commenting that:

> My own view is that in considering that limited question [the primary issue of admissibility] the judge is required to do no more than to satisfy himself that a prima facie case or originality has been made out by evidence which defines and describes the provenance and history of the recording up to the moment of production in court.[2]

1 [1972] 2 All ER 699, [1972] 1 WLR 651.
2 [1972] 2 All ER 699, [1972] 1 WLR 651, 653H.

6.162 Professor Tapper expressed the view that this exercise should be conducted first by the judge, and if, on the balance of probabilities, the judge determined the evidence could go before the jury, it would then be necessary to cover the same ground again in the same way as any other question of fact that must be decided at trial.[1] On the standard of proof to be used by the judge, O'Connor LJ indicated the criminal standard of proof is to be used in the context of handwriting,[2] and in the case of *R v Minors (Craig); R v Harper (Giselle Gaile)*,[3] Steyn J, as he then was, set out the opinion of the Court of Appeal on this matter in relation to a computer print-out:

> The course adopted by the judge in one of the two appeals before us prompts us to refer to the procedure which ought to be adopted in a case where there is a disputed issue as to the admissibility of a computer printout. It is clear that in such a case a judge ought to adopt the procedure of embarking on a trial within a trial.[4]

1 Tapper, *Computer Law*, 370; see also Rosemary Pattenden, 'Authenticating "things" in English law: principles for adducing tangible evidence in common law jury trials' (2008) 2 E & P 273–302 and 'Pre-verdict judicial fact-finding in criminal trials with juries' (2009) 29 Oxford Journal of Legal Studies 1; in the context of s 69, PACE, Professor Smith commented that during a trial within a trial, if a document is tendered by the prosecution, the standard is beyond reasonable doubt, and if tendered by the defence,

the standard is presumably on the balance of probabilities: *R v Shephard* [1993] Crim LR 295, 296.
2 *R v Ewing* [1983] QB 1039, [1983] 2 All ER 645, [1983] 3 WLR 1, CA.
3 [1989] 2 All ER 208, [1989] 1 WLR 441, CA; [1989] Crim LR 360.
4 [1989] 2 All ER 208, [1989] 1 WLR 441, 448.

6.163 He went on to indicate that the judge should apply the ordinary standard of criminal proof in reaching a decision, and in the case of *R v Neville*,[1] the members of the Court of Appeal also noted that trial judges 'should examine critically any suggestion that a prior computer malfunction has any relevance to the particular computer record tendered in evidence'.[2] The decision of the Court of Appeal in *R v Minors (Craig); R v Harper (Giselle Gaile)* to require a judge to apply the ordinary standard of criminal proof in reaching a decision when hearing evidence in a trial within a trial overrules the decision of Shaw J in *R v Robson (Bernard Jack); R v Harris (Gordon Federick)* (in which he reached an opinion that the standard was on a balance of probabilities[3]) although there is much to commend the view of Shaw J when he suggested that the prosecution need to do no more than set up a prima facie case in favour of the authenticity of the evidence:

> It may be difficult if not impossible to draw the philosophical or theoretical boundary between matters going to admissibility and matters going properly to weight and cogency; but, as I have already said, it is simple enough to make a practical demarcation and set practical limited to an inquiry as to admissibility if the correct principle is that the prosecution are required to do no more than set up a prima facie case in favour of it. If they should do so, the questioned evidence remains subject to the more stringent test the jury must apply in the context of the whole case, namely, that they must be sure of the authenticity of that evidence before they take any account of its content.[4]

1 [1991] Crim LR 288.
2 [1991] Crim LR 288, 289.
3 [1972] 1 WLR 651, 656C; this standard was agreed by counsel on both sides at 653E.
4 [1972] 1 WLR 651, 655H-656A.

6.164 The standard that a judge must apply in determining the admissibility of a videotape was considered by Cameron JA in the Canadian case of *R v Penney*[1] before the Newfoundland and Labrador Court of Appeal in 2002. In this instance, the prosecution sought to adduce evidence of the killing of marine animals. The evidence comprised of a video recording of the killing of a seal. The recording had been frequently switched on and off, as the operator of the camera selected scenes to record. The recording was filmed in mini-digital format, transferred to Beta format and then to VHS format. Before the Crown took possession of the tape, it had been in the possession of a professional editing studio for several months. There was no attempt to provide for the security or to restrict access to the tape. The Crown called the camera operator and the owner of the company for whom the camera operator worked to give evidence during the trial within a trial. The trial judge concluded that the witnesses were not credible and failed to tell the truth. He therefore refused to admit the video recording in any format. The Crown appealed to the summary appeal conviction court, which allowed the appeal. A subsequent appeal to the Newfoundland and Labrador Court of Appeal reversed the summary appeal conviction court decision and the decision of the trial judge was restored. Cameron JA addressed the issue of the standard that a trial judge should apply in determining the admissibility of videotape evidence, indicating that:

> The issue then is whether in making this finding the trial judge was usurping the role of the jury (or in this case the role of the judge at trial) or was properly carrying out the function of the judge on determination of the admissibility of hard evidence.[2]

1 (2002) 163 CCC (3d) 329.
2 (2002) 163 CCC (3d) 329, [40].

6.165 He went on:

> [43] In my view, this consideration is really a matter of weighing prejudice against probative value, in much the same way that a trial judge must examine many other kinds of evidence.
>
> [44] It is the question of fairness and absence of any intention to mislead that is really at issue in this case. The trial judge on a *voir dire* must determine whether a videotape being offered in evidence has been edited in such a way as to distort the truth.

6.166 Reference was made to *R v Nikolovski*,[1] which established that where a videotape has not been altered or changed, and where it depicts the scene of a crime, then it becomes admissible and relevant evidence.[2] In *R v Bulldog*,[3] the members of the Court of Appeal of Alberta considered this issue, and emphasized that 'What matters with a recording, then, is not whether it was altered, but rather the degree of accuracy of its representation'. In *R v Penney* the judge addressed the problem of the falsification of evidence by pointing out that the members of a jury 'can be expected to have, if not experience with, knowledge of the possibilities for manipulating the content of photographs and videotapes', and concluded that the 'standard by which the trial judge is to determine the question is on the balance of probabilities.'[2]

1 (1996) 111 CCC (3d) 403.
2 In *R. v Andalib-Goortani*, 2014 ONSC 4690 (CanLII), the prosecution failed to establish the authenticity of a digital image obtained from the Internet: the metadata had been removed, and it was not possible to ascertain the provenance of the image.
3 2015 ABCA 251 (CanLII); 326 CCC (3d) 385; [2015] AJ No 813 (QL).
4 2015 ABCA 251 (CanLII), [32].
5 (1996) 111 CCC (3d) 403, [46]–[47].

6.167 If the standard of proof of a trial within a trial is the criminal standard, it can be argued that the prosecution is required to prove its case twice: once to the trial judge and a second time before the members of the jury. Arguably, the duty of the trial judge is to sift the evidence sufficiently to establish whether the evidence is to go before the members of the jury in cases where the authenticity of the evidence is questioned by the defence.

A protocol for challenging the authenticity

6.168 Should it become the norm for the defence to challenge the authenticity of evidence in digital form, consideration, it is suggested, might be given to the development of a protocol to deal with such challenges. First, it might be necessary for the defence to warn the trial judge in advance that it will question the authenticity of identified aspects of the evidence, and to set out the grounds upon which the challenge is made.[1] Such an approach would be entirely consistent with the trial management procedures set out in Part 3, rule 3.3(2)(c)(ii) of the Criminal Procedure Rules 2015

(as amended in April 2016). If this first hurdle is overcome, then it will be for the trial judge to decide whether a trial within a trial is necessary, and if so, to set out the parameters, including the standard of proof, for which a ruling is required.

1 To a certain extent this might be already happening, for which see Oriola Sallavaci, 'Streamlined reporting of forensic evidence in England and Wales: is it the way forward?' 20 E & P 235.

6.169 As all judges are only too well aware, there is a danger that the trial judge may be seen to usurp the functions of the members of the jury in reaching preliminary decisions on authenticity when conducting a trial within a trial. Marshall J, in delivering the judgment of the Court of Appeal in the case of *R v Ali (Maqsud); R v Hussain (Ashiq)*,[1] indicated that conducting a trial within a trial should be a rare occurrence:

> In the view of this court the cases must be rare where the judge is justified in undertaking his own investigation into the weight of the evidence, which, subject to proper directions from the judge, is really the province of the jury, but the court sees that there can be cases – but they must be rare – where the issues of admissibility and weight can overlay each other.[2]

1 [1966] 1 QB 688, [1965] 2 All ER 464, [1965] 3 WLR 229, CA.
2 [1966] 1 QB 688 at 703C, [1965] 2 All ER 464, [1965] 3 WLR 229.

6.170 This restricted view was reinforced by the comments of Kilner Brown J in *R v Stevenson*:

> … as a general rule it seems to me to be highly undesirable, and indeed wrong for such an investigation to take place before the judge. If it is regarded as a general practice it would lead to the ludicrous situation that in every case where an accused person said that the prosecution evidence is fabricated the judge would be called upon to usurp the functions of the jury.[1]

1 [1971] 1 WLR 1, 4E.

6.171 Where the matter of authentication is raised, the trial judge is required to decide whether to conduct a trial within a trial. Where the decision is made to hold a trial within a trial, it will be useful for the judge to set out the scope of the hearing. In *R v Robson (Bernard Jack); R v Harris (Gordon Federick)* Shaw J said that where such a hearing takes place, it should be defined narrowly.[1] This must be right.

1 [1972] 2 All ER 699, [1972] 1 WLR 651, 655H.

6.172 In respect of the costs of such an exercise, in *R v Saward*,[1] the prosecution sought the admission of recordings of telephone conversations that were intercepted by the Dutch police and stored on a CD. The judge was invited to conduct a trial within a trial to determine whether or not the data recorded on the CD, transferred from a mainframe computer located in the Netherlands, was admissible in evidence as authentic and accurate and a reliable copy. The trial within a trial lasted for four days, and a number of witnesses, including British officers and a Dutch police officer, were called to give evidence. Lady Justice Hallett commented on the costs of such an exercise:

> Given the evidence available to the Crown we also have reservations about the profitability of the four day exercise of putting the Crown to strict proof of the exhibit. All of those involved in the conduct of criminal trials must be aware by now of the constraints upon resources and we are far from persuaded that this was a proper use of limited resources.[2]

1 [2005] EWCA Crim 3183.
2 [2005] EWCA Crim 3183, [44].

6.173 However, the defence drew a number of errors in the CD recording to the attention of the trial judge, and it was only right that this issue should be considered.

6.174 When collecting digital evidence, the investigator needs to pay careful attention to the process by which the evidence was obtained, and to demonstrate the provenance of the evidence. In *R v Skinner*,[1] the defence called into question evidence of screen images obtained by a police constable when conducting an investigation into indecent photographs of children. In the trial within a trial, the police officer gave evidence that he had a 'source' for the screen images. He admitted entering this website that he was not prepared to identify, and could only provide limited information about the provenance of the material he produced for the purposes of the investigation: namely, images that appeared on screen that were produced in the form of a print-out. He refused to name or identify the website he had entered. It was held by the members of the Court of Appeal that the trial judge wrongly admitted the evidence. First, the members of the Court accepted that it was probable that the screen images were real evidence, because their content did not require any computer input, and likened the image to somebody switching on a television set. However, the print-outs were not authenticated properly under the provisions of s 27 of the Criminal Justice Act 1988, and for that reason, the trial judge should not have admitted them. Second, there was no public interest immunity hearing to enable the judge to decide whether the prosecution need not disclose or need not give evidence as to the process by which the screen image reached the police officer, or in the absence of a proper explanation, how the screen image came to be on the police officer's computer. It was conceded that a public interest immunity hearing should have been requested, and in such circumstances, the trial judge was wrong to admit the evidence.

1 [2005] EWCA Crim 1439, [2005] All ER (D) 324 (May), [2006] Crim LR 56.

Re-introduction of the common law presumption

6.175 The Law Commission proposed the repeal of s 69 of the Police and Criminal Evidence Act 1984 and a return to the common law presumption:[1]

> 'In the absence of evidence to the contrary, the courts will presume that mechanical instruments were in order at the material time.'

1 Section 69 ceased to have any effect by s 60 of the Youth Justice and Criminal Evidence Act 1999, and s 69 was also repealed by Schedule 6; Law Commission, *Evidence in Criminal Proceedings: Hearsay and Related Topics*, 13.13; Katie Quinn, 'Computer evidence in criminal proceedings: farewell to the ill-fated s.69 of the Police and Criminal Evidence Act 1984' (2001) 5 E & P, 174–87; Amanda Hoey, 'Analysis of The Police and Criminal Evidence Act, s.69 – computer generated evidence' (1996) 1 Web JCLI.

6.176 The grounds for justification were set out in paragraphs 13.7–13.11, and are set out below with the references omitted:

> The problems with the present law
>
> 13.6 In the consultation paper we came to the conclusion that the present law was unsatisfactory, for five reasons.

> 13.7 First, section 69 fails to address the major causes of inaccuracy in computer evidence. As Professor Tapper has pointed out, 'most computer error is either immediately detectable or results from error in the data entered into the machine'.
>
> 13.8 Secondly, advances in computer technology make it increasingly difficult to comply with section 69: it is becoming 'increasingly impractical to examine (and therefore certify) all the intricacies of computer operation'. These problems existed even before networking became common.[1]
>
> 13.9 A third problem lies in the difficulties confronting the *recipient* of a computer produced document who wishes to tender it in evidence: the recipient may be in no position to satisfy the court about the operation of the computer. It may well be that the recipient's opponent is better placed to do this.
>
> 13.10 Fourthly, it is illogical that section 69 applies where the document is tendered in evidence, but not where it is used by an expert in arriving at his or her conclusions, nor where a witness uses it to refresh his or her memory. If it is safe to admit evidence which relies on and incorporates the output from the computer, it is hard to see why that output should not itself be admissible; and conversely, if it is not safe to admit the output, it can hardly be safe for a witness to rely on it.
>
> 13.11 At the time of the publication of the consultation paper there was also a problem arising from the interpretation of section 69. It was held by the Divisional Court in *McKeown v DPP* that computer evidence is inadmissible if it cannot be proved that the computer was functioning properly – even though the malfunctioning of the computer had no effect on the accuracy of the material produced. Thus, in that case, computer evidence could not be relied on because there was a malfunction in the clock part of an Intoximeter machine, although it had no effect on the accuracy of the material part of the printout (the alcohol reading). On appeal, this interpretation has now been rejected by the House of Lords: only malfunctions that affect the way in which a computer processes, stores or retrieves the information used to generate the statement are relevant to section 69.

1 It may be the case that computer technology made it increasingly difficult to comply with the provisions of s 69, but this is not an argument to presume that mechanical instruments were in order at the material time. Professor Les Hatton in his article, 'The chimera of software quality' (2007) 40 Computer 103, stated that:

> … computer programs are fundamentally unquantifiable at the present stage of knowledge, and we must consider any proof based on them flawed until we can apply the same level of verification to a program as to a theorem.
>
> Scientific papers are peer reviewed with a long-standing and highly successful system. The computer programs we use today to produce those results generally fly somewhere off the peer-review radar. Even worse, scientists will swap their programs uncritically, passing on the virus of undiscovered software faults.
>
> That the peer review process is successful is debatable – the scientific community itself has raised concerns about the various biases that afflict the selection and review processes of scientific papers and their eventual publication.

6.177 Curiously, the authors of the report did not produce any evidence to establish whether it is generally true in the absence of contrary evidence that 'mechanical instruments were in order at the material time'. There was no evidence to demonstrate that software code should benefit from this assertion. There was also no discussion of what is meant by 'in order'. This is an important issue, bearing in mind that the presumption is a presumption without the requirement of proof of a basic fact.[1] There

was a great deal of technical material in the 1970s and 1980s to demonstrate that software errors might not be obvious. Indeed, in 1986 Professor Rudolph J. Peritz noted the following (footnotes omitted):

> … [to] grant greater credibility to computerized records … because they have not been touched by 'the hand of man' succumbs to two delusions. First, it is the hands and intellects of men and women that produce computers and the programs that guide them. To believe that the absence of direct physical contact means that records are untouched betrays a naive view of electronic data processing, one that ignores the centrality of humans to any computer system's functioning. Second, trustworthiness is equated with electronic processing and opposed to human reckoning. … It ignores, for example, the great dangers of traceless change and unauthorized access, as well as the benefits of having the proponent present evidence to prove systemic accuracy.
>
> …
>
> Throughout law's intellectual history, scholars and jurists have sought methodological objectivity to justify legal decision making. … The jurisprudential lure of computer technology is a perceived absence of discretion. Once designed, built, and programmed, the machinery objectively executes the will of its creators, and thus is perceived as trustworthy. But closer scrutiny reveals, at best, a paradox of complete submission and complete autonomy. A computer performs relentlessly just as we have designed and programmed it, and in so doing, it is entirely independent of us. Computerized records also are treated as trustworthy for a second reason—because the technology is perceived as error-free. Moreover, even on those exceptional occasions of technological failure, we believe, a computer will still inform us that an error has occurred. In sum, we have come to believe that unacknowledged error and subjectivity are not only undesirable, but also indigenous to the human domain.
>
> But experience can teach us that such idealization of technology is a mirage that obfuscates the overlapping horizons of humans and computers, as well as their distinctive characteristics. In the human drama of litigation, better attention to the pragmatic jurisprudence of the Federal Rules of Evidence, as well as to the thoughtful practice recommended by the *Manual for Complex Litigation*, can help to dispel such harmful illusions. The concrete result of this attention will be the extension to the objecting party and to the court of a fair opportunity to evaluate the trustworthiness of all documents generated from computerized data.[2]

1 Katie Quinn, 'Computer evidence in criminal proceedings: farewell to the ill-fated s.69 of the Police and Criminal Evidence Act 1984' (2001) 5 E & P 174–87, 182.
2 Peritz, 'Computer data and reliability', 1001–2.

6.178 At the time of writing this article, Professor Peritz was a Visiting Associate Professor of Law at Benjamin N. Cardozo School of Law, and had worked with computers since 1962 as a programmer, operator, systems engineer, and legal consultant. He was fully conversant with the errors that occurred regularly regarding software code.

6.179 In England & Wales, section 69 was subsequently repealed,[1] and a similar reform was adopted with respect to evidence in electronic form for civil proceedings with the passing of the Civil Evidence Act 1995. It is suggested that the presumption, as set out above, that 'mechanical instruments were in order at the material time' remains far too crude an assumption to apply to computers. The authors of the Law Commission Report cite excellent reasons as to why the criminal law might be amended, but the proponents of the presumption should establish what they mean by

the term 'mechanical instruments were in order at the material time' when referring to computers or computer-like devices. A fundamental problem is caused by the fact the software errors can be present (in large numbers), but not observable in use until a specific situation is encountered.[2] For example, the 'Shellshock' vulnerability (CVE-2014-6271[3]) had been dormant in a program called Bash since 1989, which was used in Unix systems for years.

1 By s 60 of the Youth Justice and Criminal Evidence Act 1999.

2 Stephen Castell, 'Computers trusted, and found wanting' (1993) 9 Computer L & Secur Rep 155; Stephen Castell, 'Letter to the Editor' (1994) 10 Computer L & Secur Rep 158 – the views expressed by Dr Castell, despite their age, remain valid; Student Comment, 'A reconsideration of the admissibility of computer-generated evidence' (1977) 126 University of Pennsylvania Law Review 425; George L Paul, 'Systems of evidence in the age of complexity' (2014) 12 Ave Maria Law Review 173.

3 <https://cve.mitre.org/cgi-bin/cvename.cgi?name=CVE-2014-6271>.

6.180 Various challenges have been made in criminal proceedings as to the accuracy of speed measuring devices and breath analysis machines. Such devices rarely undergo a catastrophic failure, but they will drift from being accurate, which means a recalibration is necessary for time to time. Such devices continue to be the subject of challenge. This topic is not dealt with in any depth, because the aim of this chapter is to discuss the fragility of the software code in particular, although the drift or wearing out of components can of itself be a cause of software error if the software was never designed to cope with the changes that occur in such circumstances.[1] With rare exceptions, such challenges have failed. For instance, in the case of *Darby (Yvonne Beatrice) v DPP*,[2] the assertions of a police officer familiar with the use of such a device was held to be sufficient evidence to sustain the finding that the device was working correctly,[3] although where the legislation requires the date and time at which a specimen was provided to be printed on the print-out, and the date is incorrect, the machine is not considered to be capable of being 'reliable'.[4] This is supported by the comments of Kourakis and Blue JJ of the Supreme Court of South Australia in *Police v Bleeze*, who stated that 'an evidential basis for the presumption of accuracy of a scientific instrument, in a proper case, may be given by a person who, even though not a scientist with expertise in the machine's technology, is properly trained in its operation.'[5]

1 For the early history of case law, see 'The Breathalyser' by A Magistrates' Clerk (1970) 34 Journal of Criminal Law 206, and for a later analysis, see C E Bazell, 'Challenging the breathalyser' (1988) 52 Journal of Criminal Law 177 and F G Davies, 'Challenging the accuracy of the breath-test device' (1988) 52 Journal of Criminal Law 280; Ian R Coyle, David Field and Graham A Starmer, 'An inconvenient truth: legal implications of errors in breath alcohol analysis arising from statistical uncertainty' (2010) 42 Australian Journal of Forensic Sciences 101; for a discussion based on the United States of America, including an indication of the technical problems relating to fixed speed cameras, see Steven A Glazer, 'Those speed cameras are everywhere: automated speed monitoring law, enforcement, and physics in Maryland' (2012) 7 Journal of Business & Technology Law 1.

2 [1995] RTR 294, (1995) 159 J.P. 533, DC.

3 Extensive tests have indicated that many pieces of software widely used in science and engineering are not as accurate as imagined (thus affecting the accuracy of the output), and whether a police officer that has no knowledge of software code is capable of determining such a complex point is debatable: Les Hatton, 'The T experiments: errors in scientific software' (1997) 4 IEEE Computational Science & Engineering 27.

4 *Slender v Boothby* [1984] 149 J.P. 405; 'The paradox of the reliable device' (1986) 50 Journal of Criminal Law 13–15.

5 [2012] SASCF 54, [89]; for an earlier case with evidence from three witnesses, see *R v Ciantar; DPP v Ciantar* [2006] VSCA 263.

6.181 In New Zealand, Harvey J summarised the position regarding evidence of mechanical or technological devices in *R v Good*, although no evidence was proffered to substantiate the assumptions built into the presumption:

> (a) There is a presumption that mechanical instruments or technological devices function properly at the relevant time.
>
> (b) Judicial notice will be taken of the output of a notorious or well-known technology. Evidence of the way in which it works to establish that it is based on sound scientific principles is not required.
>
> (c) New or novel technologies will not receive judicial notice. Expert evidence is required to explain the operation of the technology and the scientific principles upon which it is based. Authority seems to suggest that problems have arisen when technologically based evidence has been adduced without undertaking the inquiry whether or not the technology is 'notorious' or requires expert evidence.
>
> (d) There is no rule of law which says that the reliability of the device is a precondition to admissibility. In either situation set out in (a) or (b) above the evidence is admissible – it is for the fact finder to assess weight.
>
> (e) In some cases the presumption of accuracy of a technological device will be created by statute. The manner in which the technology is operated may have an impact upon the weight to be attributed to its output.
>
> (f) In some cases devices may, as a result of their own processes, create a record which is admissible. (*R v Spiby* (1990) 91 Cr App R 186, (1991) Crim LR 199).
>
> (g) However, if there is human intervention in the performance of such processes either at the input, output or any intermediate stage, hearsay issues may arise, although in some cases exceptions to the hearsay rule may apply.
>
> (h) Whether or not there is unfairness in the process of acquiring or dealing with the evidence is a recognized common law ground to test admissibility and may be available upon the facts of each case. That is a matter primarily of human behaviour and is not intrinsically part of the technology.[1]

1 [2005] DCR 804, [70].

6.182 Proof that computers are presumed to work properly must rest with the proponent. The term 'computers' is used solely to reinforce the point that a computer or computer-like device is far more sophisticated than any pure mechanical machine, and such devices only work because a human being has written code to allow it to function. No evidence has been adduced to demonstrate the accuracy of such a presumption. One type of computer differs remarkably from another, and each will be controlled by software written by different people of varying degrees of competence to address problems of varying degrees of complexity and difficulty.[1]

1 For a discussion of software and the complex issues that affect devices used by the medical profession, see Sylvia Kierkegaard and Patrick Kierkegaard, 'Danger to public health: medical devices, toxicity, virus and fraud' (2013) 29 Computer Law and Security Review 13; Steven Hanna, Rolf Rolles, Andrés Molina-Markham, Pongsin Poosankam, Kevin Fu and Dawn Song, 'Take two software updates and see me in the morning: the case for software security evaluations of medical devices', in *Proceedings of the 2nd USENIX conference on Health security and privacy* (USENIX Association 2011).

6.183 Thus the assertion that all computers are presumed to be working properly (whatever this means) cannot be right. It is to say that all motor cars, regardless of quality, are reliable – which they demonstrably are not (although it is acknowledged that most motor cars are generally reliable). In the view of George L. Paul, 'Just

because businesses rely on faulty computer programs does not necessarily mean that courts should follow suit',[1] although in *People of the State of Colorado v Huhen* Vogt J considered that 'computer business records have a greater level of trustworthiness than an individually generated computer document'[2] without providing an authority, other than to quote from *Colorado Evidentiary Foundations*[3] that 'computers are so widely accepted and used that the proponent of computer evidence need not prove those two elements of the foundation'.

1 George L Paul, *Foundations of Digital Evidence* (American Bar Association 2008), 129; *Gordon v Thorpe* [1986] RTR 358 where two experts gave evidence of the accuracy or otherwise of a Lion Intoximeter 3000.
2 53 P.3d 735 (Colo.App. 2002), 737.
3 Roxanne Bailin, Jim England, Pat Furman and Edward J Imwinkelreid, *Colorado Evidentiary Foundations* (Michie 1997, with supplements) 736.

6.184 In this context, it is relevant to consider the decision of the Supreme Court in New Jersey in the United States, which ordered the software of a breath testing device to be reviewed in detail in the case of *State of New Jersey v Chun*.[1] In his judgment, Hoens J began by stating that 'For decades, this Court has recognized that certain breath testing devices, commonly known as breathalyzers, are scientifically reliable and accurate instruments for determining blood alcohol concentration.' This comment was based on the old technology. With the introduction of a new device, the Alcotest 7100 MK111-C, which was selected by the department of the Attorney-General, the court agreed to test the scientific validity of the machine. After extensive testing, the court concluded that the Alcotest, utilising New Jersey Firmware version 3.11, 'is generally scientifically reliable', but modifications were required to enable its results to be admitted into legal proceedings.[2] The testing of the software revealed the following issues, amongst others:[3]

> 1. That a mathematical algorithm that corrected for fuel-cell drift did not undermine the reliability of the results, but it was recommended that the machines be recalibrated every six months to ensure fuel cells are replaced regularly.
>
> 2. That a specific buffer overflow error should be corrected.
>
> 3. The court recommended that a specific number of documents be produced for the purposes of foundation of evidence.
>
> 4. That the recommendations by the defendants' experts for reorganising and simplifying the source code be considered for implementation.

1 194 N.J. 54, 943 A.2d 114.
2 943 A.2d 114, 120.
3 943 A.2d 114, 134.

6.185 The analysis of the source code indicated that there was a fault when a third breath sample was taken that could cause the reading to be incorrect, and the court saw fit to order a change in one of the formulae used in the software. Save that the extensive analysis of the device and the source code took some time and some expense, little of substance was found to be wrong with the machine. However, there are two significant points that arise as a result of this case: the first is that the software that controlled the device, written by a human, was defective, which in turn meant that the data relied upon for the truth of the statement was defective and therefore affected the accuracy and truthfulness of the evidence; and the decision by the court to intervene

by ordering certain changes and modifications to be carried out, one of which, a change in a formula, meant that part of the evidence used against drivers in the future was a set of instructions provided by the Supreme Court of New Jersey.[1]

1 There is a considerable body of case law relating to challenges of breathalyzer devices in the US. For some of the articles that discuss the position, see Workman, Jr, 'Massachusetts breath testing for alcohol'; Charles Short, 'Guilt by machine: the problem of source code discovery in Florida DUYI prosecutions' (2009) 61 Florida Law Review 177; Cheyenne L Palmer, 'DUIs and apple pie: a survey of American jurisprudence in DUI prosecutions' (2010) 13 University of the District of Columbia Law Review 407; Aurora J Wilson, 'Discovery of breathalyzer source code in DUI prosecutions' (2011) 7 Washington Journal of Law, Technology & Arts 121.

The statutory presumption

6.186 Mention might usefully be made of the powers conferred upon the Secretary of State by s 7(1)(a) of the Road Traffic Act 1988, by which a Minister may approve the use of breathalyser devices.[1] The view of the courts is illustrated in *Richardson v DPP*, in which Stanley Burnton J noted that 'The device so approved is assumed to be an effective and sufficiently accurate device for the purposes of section 5(1)(a), and that is the end of the matter.'[2] The effect is to create a statutory presumption for breathalyser devices. He went on to indicate that if the device and software approved in 1998 had changed, it was not relevant:

> On the face of it, therefore, it would seem that a device which did not include the Intoximeter EC/IR Gas Delivery System, by way of example, or the software version of which was not UK5.23, but some significantly different version, would not be an approved device. It does not follow from that that every modification to an Intoximeter takes it out of the approval. Far from it. The alteration must be such, in my judgment, that the description in the schedule to the order no longer applies to it.[3]

1 'Approval of breath test device' (1968) 32 Journal of Criminal Law 255; 'Trying times for breath testers' (1969) 33 Journal of Criminal Law 106; 'Proof of approval of "Alcotest"' (1969) 33 Journal of Criminal Law 168; 'Proof of approval by letter' (1969) 33 Journal of Criminal Law 204; 'Judicial notice of Alcotest' (1970) 34 Journal of Criminal Law, 107.
2 [2003] EWHC 359 (Admin), [6].
3 [2003] EWHC 359 (Admin), [9]; identical comments were made by Robert Goff LJ in *R v Skegness Magistrates' Court, Ex parte Cardy* (1985) RTR 49 at 61.

6.187 In *Fearnley v Director of Public Prosecutions*, Mr Justice Field observed that:

> Whilst the defence statement purports to put the prosecution specifically to proof that the software was UK 5.23, this did not mean that the prosecution had specifically to prove this matter. This is because of the general presumption that flows from the fact that the machine was of a type that had been approved,[1] this being a presumption which in my view is plainly consistent with Article 6 ECHR. Thus, it was for the appellant to adduce some evidence that the software was otherwise than the specified software before the prosecution came under a burden to prove the software. At no stage did the appellant raise or adduce such evidence and therefore he can have no substantial complaint that the prosecution were allowed to provide specific proof of the software through the engineer's report.[2]

1 Illustrating a confusion between common law and statutory presumption.
2 [2005] EWHC 1393 (Admin), [34].

6.188 In *Kemsley v DPP* Buxton LJ stated the opinion of the court on this matter:

> The statutory presumption as to approval of a particular device was conclusive as to the correctness of that device. That point does not now appear in this case, and should not appear in any case in the future.[1]

1 [2004] EWHC 278 (Admin), [11].

6.189 In *DPP v Wood; DPP v McGillicuddy*, Ouseley J indicated that if the breath test device is approved, it is therefore reliable: 'There is a common law presumption that the breath test device, if type approved, is reliable.'[1] Alternatively, where a device is weighted in favour of the accused, it is not an improper use of the device.[2] The same position is held in cases relating to speed measuring devices,[3] although if the road markings that are placed on the road to provide a scale for the digital device to measure speed are not the correct distance apart, the device will give a false reading.[4] This approach might be appropriate, given that the accused can agree to have a sample of blood sample taken, and at the same time, a copy sample of the blood is also provided to the accused. An analysis of the blood is more accurate, and the blood sample can thus be analysed by the police and independently by a person for the accused. If this option is take up by the accused, the evidence is more compelling.[5] Lord Hughes offered a further rationale in *Public Prosecution Service v McKee (Northern Ireland)*,[6] where the appellants had their fingerprints taken at the police station using an electronic device called 'Livescan'. A match was subsequently made, which the Crown relied upon at trial. Livescan devices were in general use in Northern Ireland from 2006 and throughout the period 2007–9 when statutory type approval was required by article 61(8B) of the Police and Criminal Evidence (Northern Ireland) Order 1989,[7] although approval never occurred. The appeal was dismissed. Of relevance in this context are the remarks by Lord Hughes:

> The control fingerprints taken from the appellants in the police station were not snapshots. The impressions which their fingers provided could be reproduced at any time afterwards, and would be the same. The accuracy of the Livescan readings, if disputed, could readily be checked independently by the appellants providing more samples, whether by ink and paper or by any other means, for examination by an independent expert.[8]

1 [2006] EWHC 32 (Admin), [2]; also noted by Mr Justice Cresswell in *DPP v Brown (Andrew Earle); DPP v Teixeira (Jose)* [2001] EWHC Admin 931 at [43].
2 *Ashton v DPP* (1996) 160 JP 336.
3 Section 20 of the Road Traffic Offenders Act 1988 as amended; *Griffiths v Director of Public Prosecutions* [2007] EWHC 619 (Admin), [2007] RTR 44.
4 Bill Gardner, 'Driver defeats speeding ticket with tape measure', *The Telegraph* (15 December 2014), <www.telegraph.co.uk/news/uknews/road-and-rail-transport/11294579/Driver-defeats-speeding-ticket-with-tape-measure.html>.
5 As noted by Mr Justice Newman at [8] in *The Queen on the application of Dhaliwal v Director of Public Prosecutions* [2006] EWHC 1149 (Admin); the position is similar in South Australia: *Police v Bleeze* [2012] SASCF 54, although the timing of the taking of the blood sample might be relevant, for which see *Evans v Benson* (1986) 46 SASR 317.
6 [2013] UKSC 32, [2013] 1 WLR 1611, [2014] Crim LR 77, [2013] WLR(D) 199, [2013] 3 All ER 365, [2013] 2 Cr App R 17, [2013] NI 133.
7 1989 No. 1341 (N.I. 12); article 61(8B) was repealed by the Policing and Crime Act 2009 (c. 26), ss. 112(1)(2), 116(6), Sch. 7 para. 128(2), Sch. 8 Pt. 13.
8 [2013] UKSC 32, [15].

6.190 Lord Hughes rejected the analogy between the Livescan device and speed guns and breathalysers. The latter device records an action that cannot be subsequently re-measured. Unlike a breath test, the digital data comprising the impressions of the fingerprints were reproducible, and further tests could be carried out. For this reason, it is argued, it is appropriate to expect the device to produce reliable evidence, which in turn infers that such devices have been investigated and approved by the relevant authorities.

6.191 In essence, this is what the defendants tried to achieve in *R v Skegness Magistrates' Court, Ex parte Cardy*.[1] In the absence of the right to obtain discovery as it was then called, solicitors for the accused sought to obtain relevant documents for the purpose of challenging the reliability of the Lion Intoximeter 3000 device by issuing witness summonses. Robert Goff LJ, as he then was, described the witness summonses as a means to obtain the discovery of documents, which was not permitted. Correct as this decision was, the judge commented on several occasions[2] that, in the judgment of the court, the documents that the defendants sought to obtain were not likely to be of material relevance – but failed to give any reason as to why such a conclusion was reached, given that some of the records that were requested included details of the micro-processor program and the standard operating procedures, which were highly relevant. The judge also indicated[3] that the court had been assured (it is not clear by whom) that the Home Office constantly monitored the device, and that if the devices were not reliable, the Secretary of State would not have approved their use.[4] In effect, the court was presuming the 'reliability' of such devices because the Secretary of State has so provided. Where the defence is not given the opportunity to understand how such a device is constructed, and how new versions of software affect the accuracy of the device, defendants are, it seems, not permitted to obtain any evidence to challenge the 'reliability' or 'accuracy' of the machine. The failure to provide for the proper scrutiny of digital evidence and the emphasis on relying on the assurances of the owner or user of the digital device means that the 'reliability' or 'accuracy' of these devices cannot be readily challenged in English courts.

1 [1985] RTR 49; see also *R v Coventry Magistrates' Court Ex p. Perks* [1985] RTR 74.
2 [1985] RTR 49, 57F, 57J–K, 58B–C, 58J and 59A.
3 [1985] RTR 49, 60J.
4 [1985] RTR 49, 61F–G.

Challenging the presumption

6.192 The presumption acts to place an evidential burden on the party opposing the presumption, and if he succeeds, the relying party is required to discharge the legal burden in relation to the 'reliability' of the machine, and therefore the authenticity or integrity and the trustworthiness of the evidence. The proponent must prove the authenticity of the evidence before it is admitted and can be relied upon, yet this presumption acts to bypass this requirement. It might be right to have the presumption as an aid in the authentication of the evidence – provided the basic facts are proved – as in the original formulation of the presumption, but the modern formulation of the presumption as set out by the Law Commission acts to assert that the basic facts are already proved.

6.193 It is possible to challenge the authenticity of digital evidence in a number of ways, although many reported cases appear to indicate that a lawyer will challenge the authenticity or reliability of the evidence on what might appear to be somewhat slender grounds,[1] and the judge will then have to determine whether to conduct a trial within a trial (if a criminal case) to receive evidence on the point. For instance, in *R v Coultas*,[2] the accused was convicted of dangerous driving. Evidence from the defendant's mobile telephone indicated that she was probably writing a text message when she collided and killed the cyclist. Counsel for the defendant asserted, without any foundational evidence, that there was some fault in the network coverage that would demonstrate that the defendant was probably not writing a text message at the material time. Rix LJ accepted that if such an issue had been raised at an earlier stage in the proceedings, it would have been a matter for the Crown to cover, but there was nothing about this in the defence statement, and the issue was not relevant at appeal.[3] In *The People v Lugashi*,[4] the defence argued that the prosecution had, in effect, to disprove the possibility of error before digital records of credit card fraud were admitted. Ortega J said that the 'proposed test incorrectly presumes computer data to be unreliable',[5] which does not follow. However, the appeal on this point was dismissed on a number of grounds, one of which was that the appellant did not challenge the accuracy of the information recorded in the print-out.

1 Although a letter from the defence to the prosecution putting the validity of the information of a machine in issue is not sufficient in New Zealand: *Police v Scott* 30/5/97, HC Rotorua AP89/96 – a decision that must be right and probably would be followed in other jurisdictions.
2 [2008] EWCA Crim 3261, 2008 WL 5725548.
3 [2008] EWCA Crim 3261, [21].
4 205 Cal.App.3d 632 – Ortega J reviewed relevant case law up to the date of this judgment, 27 October 1988.
5 205 Cal.App.3d 632, 640.

6.194 The problem for the lawyer making the challenge is that only the party in possession of the electronic evidence has the ability to understand fully whether the computer or computers from which the evidence was extracted can be trusted. The authors of the Law Commission paper *Evidence in Criminal Proceedings: Hearsay and Related Topics* point out that a party might rely on evidence from a computer owned or controlled by a third party that is not a party to the proceedings. However, this should not prevent the party from making the challenge of providing a suitable foundation to justify most challenges. Reed and Angel indicate that there are two broad arguments that can be pursued:

> 1. Where the party adducing the evidence does so to prove the truth of the output, it may be that the other party will challenge the accuracy of the statement by proposing that the computer, or computer-like device, exhibited faults, errors or other forms of failure that might have affected the integrity and trustworthiness of the evidence, and thus its reliability. The reliability of the computer program that generated the record may be questioned. In addition, there might be a fault with the hardware.
>
> 2. The conduct of a third party (this phrase is meant to be construed widely to include any person who does not have the authority to alter how a computer or computer-like device operates, other than the way it is intended to operate) generated the faults, errors or other forms of failure that might have affected the integrity and trustworthiness of the evidence, and thus its reliability. For instance, this can include a claim that the records were altered, manipulated, or

> damaged between the time they were created and the time they appear in court as evidence, or the identity of the author may be in dispute: the person identified as being responsible for writing a document in the form of a word processing file may dispute they wrote the text, or it might be agreed that an act was carried out and recorded, but at issue could be whether the person alleged to have used their PIN, password or clicked the 'I accept' icon was the person that actually carried out the action.[1]

1 Chris Reed and John Angel, *The Law and Regulation of Information Technology* (6th edn, Oxford University Press 2007), 596; the following analysis closely follows that of Reed and Angel, and the author is indebted to them.

6.195 The first argument was considered in the case of *DPP v McKeown; DPP v Jones*[1] over the inaccuracy of a clock in a Lion Intoximeter 3000[2] and whether the inaccuracy of the clock affected the facts relied upon as produced by the device, which was otherwise in working order. The court concluded that if there was a malfunction, it was only relevant if it affected the way in which the computer processed, stored or retrieved the information used to generate the statement tendered in evidence. This must be right.

1 [1997] 1 All ER 737, [1997] 1WLR 295, [1997] 2 Cr App R 155, HL; Philip Plowden, 'Garbage in, garbage out – the limits of s 69 of the PACE Act 1984' (1997) 61 Journal of Criminal Law 310–12; for an earlier case where the defence challenged the accuracy of the Intoximeter print-out, see *Ashton v Director of Public Prosecutions*, Times, 14 July 1995, (1996) 160 JP 336, *Ashton v DPP*, Journal of Criminal Law, (1996) 60, 350.

2 The range of approved devices constantly alters, but the case law relating to older devices remains relevant. For a more detailed discussion, see the most up-to-date edition of *Wilkinson's Road Traffic Offences* (Sweet & Maxwell).

6.196 Where the evidential burden has been successfully raised to challenge an aspect of the digital data (whether it be the integrity or reliability),[1] then the persuasive burden will on the party denying any error to prove the computer (normally the software), computer-like device or computer system is not at fault, thus demonstrating the reliability, integrity and trustworthiness and therefore the authenticity of the evidence tendered. One test is to determine how many important or critical updates of the software were made available and downloaded before the material time, and whether, if such updates were downloaded, they had a detrimental effect on the subsequent operation of the software. Claimants face a considerable problem with ATM cases, because so much can go wrong, and it can be difficult to raise sufficient evidence to shift the burden: an outsider or a bank employee might have subverted the system or a part of the system or a hardware device forming part of the ATM network (or a cloned card is used) in such a way that money is stolen from the account of an individual.[2] In such circumstances, the electronic record adduced to prove the transaction may be perfectly reliable – what will be at issue is how the thief subverted the network to steal the money. In the case of *Marac Financial Services Ltd v Stewart*,[3] Master Kennedy-Grant observed:

> The use of computers for the recording of transactions on accounts such as the cash management account in this case is sufficiently well established for there to be a presumption of fact that such computers are accurate.[4]

1 As in *Young v Flint* [1987] RTR 300, where the defence wished to cross examine the witness respecting modifications made to the device to determine whether the machine ceased to be an

approved device.
2 Ken Lindup, 'Technology and banking: lessons from the past' (2012) 9 Digital Evidence and Electronic Signature Law Review 94; Roger Porkess and Stephen Mason, 'Looking at debit and credit card fraud' (2012) 34 Teaching Statistics 87.
3 [1993] 1 NZLR 86.
4 [1993] 1 NZLR 86, [40]. Examples of where banks have not been found to be fully in control of their systems include *Patty v Commonwealth Bank of Australia* [2000] FCA 1072, Industrial Relations Court of Australia VI-2542 of 1996; *United States of America v Bonallo*, 858 F.2d 1427 (9th Cir. 1988); *Kumar v Westpac Banking Corporation* [2001] FJHC 159; *Sefo v R* [2004] TOSC 51; *R v Clarke* [2005] QCA 483.

6.197 Master Kennedy-Grant did not provide any evidence to substantiate this statement.

'Working properly'

6.198 The Law Commission made comments about the presumption at 13.14:

> Where a party sought to rely on the presumption, it would not need to lead evidence that the computer was working properly on the occasion in question unless there was evidence that it may not have been – in which case the party would have to prove that it was (beyond reasonable doubt in the case of the prosecution, and on the balance of probabilities in the case of the defence).

6.199 Three significant problems occur with the judicial comments on this topic: first, that there is no definition of what is meant by 'working properly'. A computer might be working 'properly' but not in the way an owner expects, and a third party can instruct a computer to do things that the owner neither authorizes nor is aware of. Second, it will not always be obvious whether the reliability of the evidence generated by a computer is immediately detectable without recourse to establishing whether the software code is not at fault. This is demonstrated in the Indian case of *State v Navjot Sandhu*[1] where counsel for a number of the defendants argued on appeal that the records of mobile telephone calls were not to be trusted because of duplicate entries on the relevant print-out. The witnesses were not cross-examined on this matter at trial, although Reddi J, the judge who delivered the judgment of the Supreme Court, specifically commented upon this point:

> We feel that an innocuous error in the computer recording is being magnified to discredit the entire document containing the details without any warrant. As explained by the learned counsel for the State, the computer, at the first instance, instead of recording the IMEI number of the mobile instrument, had recorded the IMEI and cell ID (location) of the person calling/called by the subscriber. The computer rectified this obvious error immediately and modified the record to show the correct details viz. the IMEI and the cell ID of the subscriber only. The document is self-explanatory of the error. … The fact that the same call has been recorded twice in the call records of the calling and called party simultaneously demonstrates beyond doubt that the correctness or genuineness of the call is beyond doubt. … Far from supporting the contention of the defence, the above facts, evidence from the perusal of the call records, would clearly show that the system was working satisfactorily and it promptly checked and rectified the mistake that occurred.[2]

1 (2005) 11 SCC 600.
2 (2005) 11 SCC 600, 152.

6.200 Many experts in software programming would be able to demonstrate that the matters recorded on a print-out do not necessarily demonstrate that the information recorded is reliable or correct and therefore to be trusted. No evidence appears to have been adduced on this point, and the judge concluded, in absence of any relevant evidence, that:

> … the printouts pertaining to the call details exhibited by the prosecution are of such regularity and continuity that it would be legitimate to draw a presumption that the system was functional and the output was produced by the computer in regular use, whether this fact was specifically deposed to by the witness or not.[1]

1 (2005) 11 SCC 600, 152.

6.201 This issue was not canvassed at trial, which arguably was a mistake, and where such evidence is to be challenged, it is necessary to lay a proper foundation. For instance, because the print-out provided information that appeared to a lay person to contain information that they might expect to see, it followed that the information recorded must have been correct. Yet there was no evidence of how faulty the software might have been at the material time (questions relating to the software up-dates might have provided some evidence to undermine the accuracy of the information contained in the print-out), nor how many subscribers challenged the accuracy of the records of telephone calls made – two items of evidence that might have raised sufficient doubts to require the prosecutor to prove the trustworthiness of the print-outs adduced at trial more fully.[1]

1 For a discussion of the evidence that can be adduced from mobile telephones and the weaknesses of such evidence, see Reg Coutts and Hugh Selby, 'Safe and unsafe use of mobile phone evidence' (Public Defenders Criminal Law Conference 2009), available online at <www.publicdefenders.nsw.gov.au/Documents/safeunsafemobilephones.pdf>; *State of Western Australia v Coates* [2007] WASC 307 (19 December 2007), cf the discussion at [211]-[220]; note also *R v Aboud; R v Stanely* [2003] QCA 499, in which location data relating to mobile telephones is considered; note the discussion of the cell site analysis in the Ontario Court of Appeal in *R. v Cyr*, 2012 CarswellOnt 16386, 2012 ONCA 919, [2012] OJ No. 6148, 104 WCB (2d) 1033, 294 CCC (3d) 421, 300 OAC 111 and in *R. v Ranger* 2010 CarswellOnt 8572, 2010 ONCA 759, [2010] OJ No. 4840, 91 WCB (2d) 271, the Ontario Court of Appeal accepted that the trial judge took judicial notice 'of the approximate location of a cell phone at the time a particular call was made based on the cell phone tower that received the signal' [14] – that is, 'He did not take judicial notice that the cell phone was in any precise location, but rather that it could properly be placed in a general location' [15].

6.202 The third problem is that the presumption asserts something positive. The opposing party is required to prove a negative in the absence of relevant evidence from the program or programs that are relied upon. In criminal proceedings, this has the unfair effect of undermining the presumption of innocence, and in civil proceedings the party challenging the presumption must convince a judge to order up the delivery of the relevant evidence, including software code, if the evidence is to be tested properly.

6.203 There is no authoritative judicial guidance in relation to the meaning of the words 'reliable', 'in order' or 'working properly' in the context of digital data. It is possible to refer to system reliability, interpreted broadly, as a measure of how a system matches the expectations of the user, but this view is problematic, because the expectations may be mistaken and can change arbitrarily, sometimes based on the user's experience. A more narrow definition is to define reliability in relation to the success with which a system provides the specified service.[1] Professor Randell and

colleagues illustrate the conundrum: 'It is of course to be hoped that the reliance placed on a system will be commensurate with its reliability.' Herein lies the rub: 'Notions of reliance, therefore, can be as much bound up with psychological attitudes as with formal decisions regarding the requirement that a system is supposed to satisfy.'[2] The authors continue:

> In fact, the history of the development of computers has seen some fascinating interplay between reliance and reliability. The reliability of early computers caused relatively little reliance to be placed on the validity of their outputs, at least until appropriate checks had been performed. Even less reliance was placed on the continuity of their operation – lengthy and frequent periods of downtime were expected and tolerated. As reliability increased so did reliance, sometimes in fact outdistancing reliability so that additional efforts had to be made to reach previously unattained reliability levels. During this time computing systems were growing in size and functional capacity so that, although component reliability was being improved, the very complexity of systems was becoming a possible cause of unreliability, as well as a cause of misunderstandings between users and designers about system specification.[3]

1 Randell, Lee and Treaven, 'Reliability issues in computing system design' 123.
2 Randell, Lee and Treaven, 'Reliability issues in computing system design' 124.
3 Randell, Lee and Treaven, 'Reliability issues in computing system design' 124. That IT projects invariably cost more than estimated, overrun, and sometimes fail to be implemented is a notorious fact. A citation (or citations) is not necessary.

6.204 In considering a number of examples of reliability issues, Professor Randell indicates that the design of software is inextricably intertwined with the other factors that are responsible for the failure of computer projects:[1]

> … reliability is a commodity whose provision involves costs, either direct, or arising from performance degradation. In theory, the design of any nontrivial computing system should involve careful calculations of trade-offs between reliability, performance, and cost. In practice the data and relationships which would be needed for such calculations in complex systems, are quite often unknown, particularly with regard to unreliability caused by residual design faults.[2]

1 For a more detailed treatment of the causes of the failure of projects, see Robert L Glass, *Software Runways: Lessons Learned from Massive Software Project Failures* (Prentice Hall 1998); 'Report of the Defense Science Board Task Force on Defense Software' (November 2000), available at <www.acq.osd.mil/dsb/reports2000s.htm>; Planning Report 02-3 The Economic Impacts of Inadequate Infrastructure for Software Testing, prepared by RTI for the National Institute of Standards & Technology (May 2002), available at <www.nist.gov/sites/default/files/documents/director/planning/report02-3.pdf>; Robert N Charette, 'Why software fails' (2005) 42 IEEE Spectrum 42, available at <http://spectrum.ieee.org/computing/software/why-software-fails>.
2 Randell, Lee and Treaven, 'Reliability issues in computing system design,127.

6.205 Linden pointed out that reliability '… means not freedom from errors and faults, but tolerance against them. Software need not be correct to be reliable',[1] and Denning indicated that although '… reliability, in the sense of error tolerance, has long been sought in operating system software, it has always been difficult to achieve.'[2] Responsible practice will often include processes such as the maintenance and review of defect records, and testing or re-qualification of an upgrade before it is distributed: these are some of the issues about which questions can be legitimately asked by a party in seeking to question the presumption of 'reliability'.

1 Linden, 'Operating system structures to support security and reliable software' 361.
2 Denning, 'Fault tolerant operating systems' 359.

Concluding remarks

6.206 It is proposed that the proponents of a presumption that computers and computer systems 'were in order at the material time' should state what is meant by such a proposition if it is to remain. In *Holt v Auckland City Council*, Richardson J observed the need to provide evidence to justify reliance:

> The results depend on the manner in which it is programmed. And there is no basis on which the Court could take judicial notice of the manner in which this equipment was programmed and maintained. Evidence was necessary to justify reliance on the computer print out …[1]

1 [1980] 2 NZLR 124, 128 (35 – 40).

6.207 It does not appear that any thought has been given to demonstrating what the proposition means. The Law Commissioners specifically commented on the contrary argument made by David Ormerod, now Professor Ormerod, to their proposal to repeal s 69. Professor Ormerod 'contended that the common law presumption of regularity may not extend to cases in which computer evidence is central'.[1] This comment by Professor Ormerod must be right.

1 At 13.16.

6.208 In *Scott v Baker*,[1] Lord Parker CJ and his brother judges rejected the argument of the prosecution that there was a presumption that where an alcohol measuring device was used by the police, it therefore followed that the device was approved by the Secretary of State. The Law Commissioners agreed that this presumption must have been applicable to the Intoximeter cases, and yet noted that this had not been raised in previous cases. They then went on, at 13.17, to state (footnote omitted):

> It should also be noted that *Dillon* was concerned not with the presumption regarding machines but with the presumption of the regularity of official action. This latter presumption was the analogy on which the presumption for machines was originally based; but it is not a particularly close analogy, and the two presumptions are now clearly distinct.

1 [1969] 1 QB 659; 'Divisional court cases breath tests: approval of device *Scott v. Baker*' (1968) 32 Journal of Criminal Law 151.

6.209 Professor Ormerod referred to *Dillon* for the point that the prosecution is not entitled to rely on a presumption to establish facts central to an offence, and it is essential for the prosecution to prove, on the facts of *Dillon*, the lawfulness of the prisoner's detention by affirmative evidence.[1] In his article, Professor Ormerod argued that where evidence in digital form is fundamental, such as in bank frauds, it will be necessary to require specific proof of reliability. This proposition must be correct: the presumption on its own cannot bear the weight of proof beyond reasonable doubt.

1 David Ormerod, 'Proposals for the admissibility of computer evidence' (1995) 6 Computers and Law 24.

6.210 In the absence of evidence that such a presumption can possibly apply to such complex objects as computers and computer systems, it is suggested that any presumption that a computer or computer-like machine is working properly be guided by considerations as to how 'correct operation', 'quality', 'reliability' and 'integrity' can be incorporated within the evaluation of the presumption.[1] It cannot be right to infer 'reliability' from reliance.

1 The Commonwealth Draft Model Law on Electronic Evidence also refers to 'reliability', and the expert group noted that 'The Group agreed that system reliability is the most sensible measurement', for which see the Model Law (replicated in full in appendix 1), para 2 of the introductory remarks.

6.211 As it stands, the presumption places an evidential burden on the party opposing the presumption, described by Tipping J: 'The accused must be able to point to a sufficient evidential foundation for the suggestion that the device was unreliable in the relevant sense, before being entitled to have the point considered by the jury',[1] and it may be that careful consideration ought to be given to the hurdle a party must overcome in order to meet the evidential burden. In this respect, the defence was correct to challenge the evidence of the CD which contained the intercepted recordings in *R v Saward*,[2] because had the prosecution more thoroughly ensured the continuity of the evidence, it is possible the defence may not have had a legitimate objection. In *Scott v Otago Regional Council*, Heath J indicated that cross-examination of relevant points can be sufficient to put the point in issue, which must be right (although the cross examination might more usefully have also considered questioning how many software updates were provided by the manufacturer of the product that corrected faults):

> No evidence was offered about the reliability of the computer and software used to establish that they were 'of a kind that ordinarily [do] what a party asserts [them] to have done'.[3] Mr Reeves offered no evidence that he had used the programme successfully in the past and had found it to be working normally. Nor was there any independent evidence to explain how the computer programme worked and what it could reliably be expected to do. In a prosecution such as this, Mr Andersen's cross-examination of Mr Reeves was sufficient to put the point in issue.[4]

1 *R v Livingstone* [2001] 1 NZLR 167, [5].
2 [2005] EWCA Crim 3183.
3 Where the basic fact of the presumption is not satisfied, the presumption fails.
4 CRI 2008-412-17-20, High Court Dunedin, 3 November 2008, [2008] Your Environment 392; 31 TCL 48/8, [33].

6.212 The Law Commission indirectly discussed 'reliability' at para 13.18, but only by referring to the possibility of a 'malfunction'. The entire discussion seems to be predicated upon machines used to test the amount of alcohol a person has consumed, rather than the very much broader range of computers and computer-like devices that are in common use:

> Even where the presumption applies, it ceases to have any effect once evidence of malfunction has been adduced. The question is, what sort of evidence must the defence adduce, and how realistic is it to suppose that the defence will be able to adduce it without any knowledge of the working of the machine? On the one hand the concept of the evidential burden is a flexible one: a party cannot be required to produce more by way of evidence than one in his or her position could be

> expected to produce. It could therefore take very little for the presumption to be rebutted, if the party against whom the evidence was adduced could not be expected to produce more.

6.213 The comments by Lord Hoffmann in *DPP v McKeown; DPP v Jones*,[1] in which he offered the opinion that 'It is notorious that one needs no expertise in electronics to be able to know whether a computer is working properly',[2] can be considered to be the extreme view that will not be shared by many computer experts – or indeed lay people. His comment is not merely extreme but vacuous. It is like saying that you do not need to know the chemistry of ink to know whether writing works. This is not relevant, because you can still write nonsense, regardless of the chemical properties of the ink.

1 [1997] 1 All ER 737, 743, [1997] 1 WLR 295, [1997] 2 Cr App R 155, HL.
2 [1997] 1 All ER 737, 743b.

6.214 It is noticeable that paragraph 432 of the Explanatory Notes to the Criminal Justice Act 2003 indicated that, in respect of testimony under s 129(1):

> This section provides where a statement generated by a machine is based on information implanted into the machine by a human, the output of the device will only be admissible where it is proved that the information was accurate.

6.215 Here the emphasis is on the accuracy of the information as an input to the computer, not whether the computer was working consistently, or to put it another way, whether the system was not working in accordance with an expectation, or the ability of the computer to return generally verifiably correct results.

6.216 The problem is that Lord Hoffmann considered the issue from the opposite perspective: an assumption that the computer is working properly because of what the user can see, not what an unknown third party does not want them to see, or prevents them from seeing and understanding what else the computer is doing without the knowledge of the owner. It is debatable whether a computer operating with such a 'parasite' within its system can be considered to be 'working properly', although it is conceivable to consider a properly working computer in parallel: where a computer will operate properly in accordance with the requirements of the owner, where the computer provides verifiably correct results, while simultaneously undertaking unrelated tasks for an unknown third party, and where neither activity will impinge on the accuracy of the other.

6.217 As a matter of admissibility, it is necessary that proof that a computer, computer-like device or network (comprising many computers and modes of communication) was 'in order' at the material time – indeed, in the UK, s 129(2) of the Criminal Justice Act 2003 preserves the common law position:[1]

> 129 Representations other than by a person
>
> (1) Where a representation of any fact—
>
> (a) is made otherwise than by a person, but
>
> (b) depends for its accuracy on information supplied (directly or indirectly) by a person,
>
> the representation is not admissible in criminal proceedings as evidence of the fact unless it is proved that the information was accurate.

> (2) Subsection (1) does not affect the operation of the presumption that a mechanical device has been properly set or calibrated.

1 *Evidence in Criminal Proceedings: Hearsay and Related Topics* (Law Com no 245) (19 June 1997), 7.50.

6.218 That software is notorious for being the subject of defects leads to a somewhat uneasy state of affairs. It cannot be right to presume that a machine (in particular a computer, computer-like device or network) was 'in order' (whatever that means) or 'reliable' at the material time. The proponents of the presumption have not provided any evidence to demonstrate the accuracy of the assertion. Evidence in digital form is not immune from being affected by the faults in software written by human beings. The use of the words 'operating properly' illustrates the misconceptions described in this chapter.

6.219 The lack of any evidence to support the proposition is especially relevant in the light of the underlying rationale of evidence. In *A Philosophy of Evidence Law: Justice in the Search for Truth*,[1] Professor Hock Lai Ho demonstrates that the finder of fact acts as a moral agent, and central to this is that the findings by a court must be justifiable, and meet the demands of rationality and ethics.[2] When read in the light of the unique characteristics of digital evidence, the rationale of the evidential process takes on an even more relevant role – a role that the author might not have contemplated. This is because the factors and subsequent analysis have an added poignancy when taking into account the complexity of digital evidence: the potential volumes of evidence, the difficulty of finding evidence, persuading the judge to order additional searches or to order the disclosure of relevant digital data, the ease by which digital evidence can be destroyed, the costs of such exercises, the lawyer's lack of knowledge when dealing with this form of evidence and the presumption that computers are 'reliable' or 'working properly'. In this respect, the inadequacy of the procedure leading to trial brought about by an incomplete understanding and application of the presumption may cause unfairness.

1 Oxford University Press 2008.

2 Note the article by Louis Kaplow, 'Burden of proof' (2012) 121 Yale Law Journal 738, in which the author considers how robust the evidence ought to be in order to assign liability when the objective is to maximize social welfare.

6.220 The question is whether the presumption is to remain. The failure of the proponents to provide evidence that the presumption has any basis in fact is a strong indication that the presumption does not merit being in place – and any argument in favour of the proposition ought to clearly indicate why banking systems, manufacturers of motor vehicles, aircraft and medical devices – to name but a few – should be rewarded by such a presumption. In addition, the innumerable examples of the failure of software outlined in this chapter, and other failures that are constantly brought to our attention regularly by the media, as well as the failures we witness ourselves in our everyday lives, act to challenge why software code should benefit from such a presumption. This is particularly so when evidence in digital form is more likely to be open to challenge, as illustrated above.

6.221 In addition, considering that the presumption is only an evidential presumption, the bar for raising doubts about the reliability or otherwise of a computer, computer-

like device or network must not be placed too high.[1] For instance, in *DPP v Wood; DPP v McGillicuddy*, Ouseley J indicated (in respect of the Intoximeter EC/IR):

> The nature and degree of an alleged unreliability has to be such that it might be able to throw doubt on the excess in the reading to such an extent that the level of alcohol in the breath might have been below the level at which a prosecution would have been instituted.[2]

1 Sergey Bratus, Ashlyn Lembree and Anna Shubina, 'Software on the witness stand: what should it take for us to trust it?', in Alessandro Acquisti, Sean Smith and Ahmad-Reza Sadeghi (eds), *Trust and Trustworthy Computing: Proceedings of the Third International Conference*, TRUST 2010, Berlin, Germany, June 21-23, 2010 (Springer 2010), pp. 396–416.
2 [2006] EWHC 32 (Admin), [36].

6.222 However, as indicated by Eric Van Buskirk and Vincent T. Liu:

> The Presumption of Reliability is difficult to rebut. Unless specific evidence is offered to show that the particular code at issue has demonstrable defects that are directly relevant to the evidence being offered up for admission, most courts will faithfully maintain the Presumption of Reliability. But because most code is closed source and heavily guarded, a party cannot audit it to review its quality. At the same time, however, source code audits are perhaps the best single way to discover defects.
>
> This difficulty gives rise to an important question: if a party cannot gain access to source code without evidence of a defect, but cannot get evidence of a defect without access to the source code, how is a party to rebut the Presumption? Rather than wrestle with, or even acknowledge, this conundrum, most courts simply presume that all code is reliable without sufficient analysis.[1] [Footnotes omitted]

1 Eric Van Buskirk and Vincent T Liu, 'Digital evidence: challenging the presumption of reliability' (2006) 1 Journal of Digital Forensic Practice 20.

6.223 This view is illustrated in the case of *State of Florida v Bastos*,[1] an appeal before the District Court of Appeal of Florida, Third District, where Cope J held that source code for an Intoxilyzer 5000 breath test machine used in the defendants' cases was not 'material' within the meaning of the provisions of the uniform law to secure the attendance of witnesses from within or without a state in criminal proceedings. The judge went on to say:

> However, we cannot accept the proposition that simply because a piece of testing equipment is used in a criminal case, it follows that the source code for its computer must be turned over. There would need to be a particularized showing demonstrating that observed discrepancies in the operation of the machine necessitate access to the source code. We are unable to see that any such evidence was brought forth in the evidentiary hearing below.[2]

1 985 So.2d 37 (Fla.App. 3 Dist. 2008). In *State of North Carolina v Marino*, 747 S.E.2d 633 (N.C.App. 2013), the court refused to accept that the decisions of the Supreme Court in *Crawford v Washington*, 541 U.S. 36, 51, 124 S.Ct. 1354, 158 L.Ed.2d 177, 192 (2004), nor that the decision in *Melendez–Diaz v Massachusetts*, 557 U.S. 305, 310–11, 129 S.Ct. 2527, 174 L.Ed.2d 314, 321–22 (2009) stood for the proposition that a defendant had a right under the Sixth Amendment to examine the Intoximeter source code. But see *In re Commissioner of Public Safety v Underdahl*, 735 N.W.2d 706 (Minn. 2007) and *State of Minnesota v Underdahl*, 767 N.W.2d 677 (Minn. 2009), where it was held that an order that the Commissioner of Public Safety provide Mr Underdahl with an operational Intoxilyzer 5000EN instrument and the complete computer source code for the operation of the device was affirmed partly on the basis that the State had possession or control of computer source code for the purposes of discovery.

2 985 So.2d 37 (Fla.App. 3 Dist. 2008), 43.

6.224 The party contesting the presumption will rarely be in a position to offer substantial evidence to substantiate any challenge[1] because the party facing the challenge will generally (but not always) be in full control of the computer or computer systems that are the subject of the challenge.[2] Offering an explanation that is not reinforced with any evidence will not be sufficient, for which see *Burcham v Expedia*, Inc.,[3] and a theory that is 'incredible' should not require the court to consider the matter in any detail.[4] The lack of evidence to rebut the presumption is not helpful, for which see *Public Prosecution Service v McGowan*.[5] From the perspective of criminal procedure, it must be right that the defence should give the prosecution advance notice that they intend to challenge the device, as suggested by Newman J in *Director of Public Prosecutions v Spurrier*:

> As a matter of general rule, I can see no reason why the defence should not be taken to be required, of course on pain of paying the costs of an adjournment if that proves to be necessary, to give some notice in advance of the trial of the grounds upon which a claim that the device was defective will be advanced.[6]

1 For an interesting discussion that includes the burden in the context of authentication, see Rudolph J Peritz, 'Computer Data and Reliability: A Call for Authentication of Business Records Under the Federal Rules of Evidence', 965–1002.

2 It is becoming increasingly common for organisations and individuals to rely on third parties to provide computing facilities, through what is termed 'Cloud computing' by the technical community; for a detailed explanation, see Stephen Mason and Esther George, 'Digital evidence and "cloud" computing' (2011) 27 Computer Law & Security Review 524.

3 2009 WL 586513.

4 For which see *Novak d/b/a PetsWarehouse.com v Tucows, Inc.*, 73 Fed. R. Evid. Serv. 331, 2007 WL 922306 affirmed *Novak v Tucows, Inc.*, 330 Fed.Appx. 204, 2009 WL 1262947.

5 [2008] NICA 13, [2009] N.I. 1.

6 [2000] RTR 60, 68 item (6).

6.225 The evidence of relevant audits is also of significance, such as where John Rusnak forged trades in a word document and an audit failed to indicate the forgery;[1] and where Nick Leeson forged data that was not noticed by audits.[2] The importance of audits was glaring revealed in *A and others (Human Fertilisation and Embryology Act 2008)*.[3] Following Cobb J's judgment in *AB v CD*,[4] the HFEA required all 109 licensed clinics to carry out an audit of their records. It transpired that 51 clinics (46 per cent) had discovered 'anomalies' in their records, including missing forms from the records, forms completed or dated after treatment had begun, incorrectly completed, unsigned, and not fully completed forms, forms with missing pages, and even forms completed by wrong persons.[5] Sir James Munby, President of the Family Division, had this to say:

> The picture thus revealed … is alarming and shocking. This is, for very good reason, a medical sector which is subject to detailed statutory regulation and the oversight of a statutory regulator – the HFEA. The lamentable shortcomings in one clinic identified by Cobb J, which now have to be considered in the light of the deeply troubling picture revealed by the HFEA audit and by the facts of the cases before me, are, or should be, matters of great public concern. The picture revealed is one of what I do not shrink from describing as widespread incompetence across the sector on a scale which must raise questions as to the adequacy if not of the HFEA's regulation then of the extent of its regulatory powers.[6]

1 Siobhán Creaton and Conor O'Clery, *Panic at the Bank: How John Rusnak Lost AIB $691,000,000* (Gill & Macmillan 2002) 96–7.
2 Nick Leeson with Edward Whitley, *Rogue Trader* (Sphere 2013), 117, 120–1, 239; see also the *Report of the Board of Banking Supervision Inquiry into the circumstances of the collapse of Barings* (ordered by The House of Commons to be printed 18 July 1995) (HMSO, 18 July 1995), chapters 9 and 10 and conclusions 13.4(b) and (c) at 232.
3 [2015] EWHC 2602 (Fam).
4 [2013] 2 FLR 1357, [2013] EWHC 1418 (Fam).
5 *A and others (Human Fertilisation And Embryology Act 2008)* 2015] EWHC 2602 (Fam), Sir James Munby P at [7].
6 [2015] EWHC 2602 (Fam), [8].

6.226 The banking cases also illustrate the nature of the problem,[1] as do the unintended acceleration cases. Crucially, in the US *Bookout* case, which was one of the high profile unintended acceleration cases, Selna J ordered the disclosure of the software code.[2] The explanation might be because of two significant, and rather fortuitous, factors. When Jean Bookout was driving her 2005 Toyota Camry, it suddenly accelerated. She took action by pulling the parking brake. By so doing, the right rear tyre left a 100-foot skid mark, and the left tyre left a 50-foot skid. The vehicle continued to speed down a ramp, across the road, and came to rest with its nose in an embankment, injuring her and killing her passenger and best friend Barbara Schwarz. Before she died, Schwarz called her husband and said 'Jean couldn't get her car stopped. The car ran away with us. There's something wrong with the car.'[3] Both the skid marks and the telephone call by Barbara Schwarz undermined any suggestion that the acceleration was due to a physical problem in the cabin of the vehicle.

1 Gerwin Haybäck, 'Civil law liability for unauthorized withdrawals at ATMs in Germany' (2009) 6 Digital Evidence and Electronic Signature Law Review 57; Stephen Mason, 'Debit cards, ATMs and negligence of the bank and customer' (2012) 27 Butterworths Journal of International Banking and Financial Law 163; M. Silalahi Nuth, 'Unauthorized use of bank cards with or without the PIN: a lost case for the customer?' (2012) 9 Digital Evidence and Electronic Signature Law Review 95; Stephen Mason, 'Electronic banking and how courts approach the evidence' (2013) 29 CTLR 144.
2 'Stipulated Protective Order' dated 16 November 2011 – R Graham Esdale, Jr, a principal at Beasley Allen, Montgomery, Alabama, USA kindly provided a copy to the author.
3 Anthony Anderson, 'Sudden acceleration, spaghetti software and trauma at the kitchen sink', *Expert Witness Journal* (Spring 2014) (no pagination) available at <http://blog.copernicustechnology.com/wp-content/uploads/2014/05/Uncommanded-Acceleration-article.pdf>; 'Sudden unintended acceleration redux: the unresolved issue', (2009) 6 The Safety Record, available at <www.safetyresearch.net/blog/articles/sudden-unintended-acceleration>; given that the *Bookout* case demonstrated the claims of the plaintiff, the decision of Carr J to exclude a number of important expert witness, while permitting the expert witness for Ford (an employee) to give evidence is to be questioned in *Buck v Ford Motor Company*, 810 F.Supp.2d 815 (N.D.Ohio 2011).

6.227 As Professor Peritz pointed out in 1986 (as did Lynda Crowley-Smith in 1996):[1]

> Computers provide an illusory basis for shortcircuiting traditional legal processes because they cannot be isolated from the people that build and run them. They simply cannot guarantee error-free processing.[2]

1 Lynda Crowley-Smith, 'The *Evidence Act* 1995 (Cth): should computer data be presumed accurate?' (1996) 22 Monash University Law Review 166.
2 Peritz, 'Computer data and reliability', 1000.

6.228 This is why lawyers and members of the judiciary need to understand two significant issues about the world in which we live now, and the reliance on modern

technology. First, the evidential presumption, which is a delusion, that software code is 'reliable' must be reconsidered. The rationale used by judges that software code is part of a 'notorious' class of machines, or the operation of computers and other such devices are 'common knowledge' must be reversed. In his speech *Science and Law: Contrasts and Cooperation* before the Royal Society in London on 25 November 2015,[1] Lord Neuberger said that 'scientists and lawyers each search for and assess hard facts from which they can establish the truth',[2] yet lawyers and judges rely on 'common sense' when many 'well-established principles are positively contrary to common sense'.[3] Justifications around loose notions of 'notorious' or 'common knowledge' in respect of software programs is irrational. Justice should not be based on concepts with no basis in logic or science. It is necessary for lawyers and judges to take account of this element of irrationality that has been the law for far too long. To resolve the problem expeditiously, an appellate court could adjust the presumption by restricting it to mechanical instruments and instruments for which statutory presumptions exist. Thereafter, it will be for the proponent to provide for the reliability (if 'reliability' is to be used) of the software. Evidence of reliability will not always be required. No doubt suitable procedural mechanisms can be put in place to allow a party to require relevant evidence of reliability where it is challenged.

1 <www.supremecourt.uk/docs/speech-151124.pdf>.
2 Lord Neuberger, *Science and Law: Contrasts and Cooperation*, [9].
3 Lord Neuberger, *Science and Law: Contrasts and Cooperation*, [13].

6.229 Second, judges should understand the necessity of requiring the disclosure of software code and relevant audits of systems, and determine whether security standards, if applied, have been applied properly.[1] These steps ensure that the judicial process more fully comprehends the evidential reality of software code and 'digital systems', and helps to preserve fairness in legal proceedings.[2]

1 Failures in banking systems used by millions of customers are demonstrated in this article: Murdoch, Bond and Anderson, 'How certification systems fail' 40.
2 Colin Tapper, 'Judicial attitudes, aptitudes and abilities in the field of high technology' (1989) 15 Monash University Law Review 219, 228, where Professor Tapper considers the members of the House of Lords and Court of Appeal were unduly restrictive regarding the transient storage of a false password in *R v Gold and Schifreen* [1989] QB 1116 (CA), [1988] 2 All ER 186 (HL).

7

Authenticating electronic evidence

Stephen Mason and Allison Stanfield

7.1 The term 'trustworthiness' is often used to describe that a thing deserves, or is entitled to, trust or confidence. There are two qualitative dimensions to the concept of trustworthiness: reliability and authenticity. Reliability is meant to demonstrate that the record is capable of standing for the facts to which it attests. Authenticity means the record is what it claims to be.[1] For evidence to be authentic, it must be proved that it is what it purports to be. It follows that it is a condition precedent to admissibility.[2] The term 'authentic' is used to describe whether a document or data are genuine, or that the document (in the case of digital data) 'matches the claims made about it'.[3]

1 Heather MacNeil, *Trusting Records: Legal, Historical and Diplomatic Perspectives* (Kluwer Academic Publishers 2000) xi; Livia Iacovino, *Recordkeeping, Ethics and Law* (Springer 2006) 41, for further comments about 'trustworthiness'.

2 Daniel K B Seng, 'Computer output as evidence' [1997] Sing JLS 161–3.

3 Rosemary Pattenden, 'Authenticating "things" in English law: Principles for adducing tangible evidence in common law jury trials' (2008) 12 E & P 275.

7.2 For a physical document, its authenticity comprises such attributes as the state of being faithful to an original, uncorrupted and with a verified provenance (encompassing the following attributes: uniqueness, unambiguity, conciseness, repeatability and comprehensibility).[1] Although electronic evidence has very different characteristics to paper, the rules of evidence that have developed with respect to the authentication of evidence, particularly documentary evidence, are still highly pertinent to electronic evidence.

1 Attributes suggested by Philip Turner, 'Digital provenance – interpretation, verification and corroboration' (2005) 2 Digital Investigation 45–9.

7.3 Yet, with its unique characteristics, complex questions about the integrity and security of electronic evidence are raised which must be examined when considering the authentication of electronic evidence. This is because, as noted by Steven W. Teppler:

> Digital data is inherently malleable or mutable. The inherently mutable nature of computer-generated data creates new issues that have a significant and detrimental effect on reliability, authentication, and ultimately on the issue of admissibility. This mutability, in turn, exposes the inherent frailty of digital data sought to be introduced as evidence.[1] (footnotes omitted)

1 Steven W Teppler, 'Testable Reliability: A Modernized Approach to ESI Admissibility' (2014) 12 Ave Maria L Rev 213, 217.

7.4 Each case is necessarily considered on its merits, and in the case of authenticating electronic evidence, there is very little clear guidance on how to determine authenticity, since traditional rules look at individual documents rather than the digital system in which digital data are created. One possible exception is Canada, where the provisions

Stephen Mason and Allison Stanfield, 'Authenticating electronic evidence', in Stephen Mason and Daniel Seng (eds.), *Electronic Evidence* (4th edn, University of London 2017) 193–260.

of the Canadian Evidence Act recognize that electronic documents are part of a computer system.[1] In addition, although the rules for authentication of evidence vary from jurisdiction to jurisdiction, all jurisdictions require that a document tendered into evidence be what it purports to be. With electronic evidence, this can be difficult to establish, especially when the lawyers presenting the evidence do not understand the nature of electronic evidence, and both the common law rules and legislative provisions embody guidelines that were developed around paper documents.

1 Chasse suggests that 'an electronic record unlike a pre-electronic paper record, is dependent upon its ERMS [Electronic Records Management System] for everything, including its existence, its accessibility and its integrity': Ken Chasse, *Guilt by Mobile Phone Tracking Shouldn't Make 'Evidence to the Contrary' Impossible* <www.slaw.ca/2016/10/04/guilt-by-mobile-phone-tracking-shouldnt-make-evidence-to-the-contrary-impossible/>.

General considerations relating to authenticity

7.5 A great deal of work has been undertaken in understanding the nature of a digital record and the dynamics of ensuring trustworthiness, reliability, integrity and authenticity,[1] and there is a link between the requirements of archivists, digital evidence professionals and lawyers.[2] Some of the work has led to the implementation of standards, such as the Electronic Records Management Software Applications Design Criteria Standard (DoD 5015.2-STD) (25 April 2007) issued by the Assistant Secretary of Defense for Command, Control, Communication and Intelligence in the US, replacing two earlier standards. Another is the protocol for the management and storage of electronically stored information (CAN/CGSB-72.11-93) issued by the Canadian General Standards Board, which was relied upon in a challenge to the admissibility of digital records in *R. v Oler*.[3] One practitioner claims that, given the change in the way digital data is now structured, 'all but a very few of them [as in protocol/standards] are either irrelevant, obsolete, over-engineered or completely unnecessary for managing the final stages of the information lifecycle of unstructured electronic content given today's technology'.[4] This may be an overstatement, as protocols and standards remain of use to lawyers when the authenticity of a document in digital form is tested in court. In summary, a project by the University of British Columbia has identified eight components of an electronic record for which issues of authenticity may arise:

> 1. The medium, being the storage medium upon which the data are stored.
>
> 2. Content, referring to the message contained in the document, which may only be made manifest when assembled into a complete whole either by rendering the text on a screen or as a print-out. When not in use, the content consists of pointers to data that will be located in different places within a database or a series of databases.
>
> 3. The physical form, comprising a range of elements, such as the script (typeface of font, formatting, inserts, the use of colours, and such like); the language used; additional information, such as attachments, comments, time-tamps; and the configuration and architecture of the operating system and records. Any change in these elements will affect the data, and may, in turn, create a different record.
>
> 4. Intellectual form comprises three elements: 'information configuration', referring to how the content is represented (text, graphics, images, sounds or any combination); 'content articulation', comprising date, salutation, exposition and 'annotations', referring to any additional information added to the record in

the execution of the document (such as the authentication of a digital signature), how the matter is handled (whether it has been designated as being urgent, together with the date and time of any action), any developments that have occurred (evidence of subsequent actions taken), and the management of the record, such as the classification of the document and its registry number.

5. Action comprises the record of the act or action that gave rise to the record. A geographical information system has the capacity to present data in different formats, perhaps geographically within and between departments.

6. People are the agents who can create the content of the data. Those responsible for the existence of the data include the author and addressee. The retention of this information is necessary for the preservation of the provenance of the data over time.

7. The archival bond in the digital environment requires a link to be created and maintained when data is removed from a system and put into an archive. It is necessary to make explicit the relationship between the records and the actions of archiving to demonstrate that the original data has not been altered.

8. The context refers to the framework of the document. Four elements have been identified: the legal and administrative context (the legal and organizational status of the body, such as body corporate, human resource); the provenance (the body that is responsible for creating the data, structure, functions); the procedural context (the procedures by which the record is created), and the documentary context (the internal structure of the archive), which in turn represents the totality of the bonds that make up the context. In the digital environment, the technological context is crucial, and is discussed more fully below.[5]

1 Charles M Dollar, *Authentic Electronic Records: Strategies for Long-Term Access* (Cohasset Associates, Inc. 2002); Luciana Duranti, Terry Eastwood and Heather MacNeil, *Preservation of the Integrity of Electronic Records* (Springer 2003).

2 Matthew G Kirschenbaum, Richard Ovenden and Gabriela Redwine, *Digital Forensics and Born-Digital Content in Cultural Heritage Collections* (CLIR Publication No. 149 2010), Council on Library and Information Resources, Washington, D.C., 1.1, available at <www.clir.org/pubs/reports/pub149/pub149.pdf>.

3 2014 ABPC 130.

4 For a critique by Don Leuders, who originally supported the standard, see 'On Why I No Longer Support the DoD 5015.2 Standard', 27 May 2013, at <http://community.aiim.org/blogs/don-lueders%20crm%20cdia/2013/05/27/on-why-i-no-longer-support-the-dod-5015.2-standard>.

5 MacNeil, *Trusting Records*, 90–96; Iacovino, *Recordkeeping, Ethics and Law*, 46–55; see also the 13 properties identified in the University of Pittsburg project: David Bearman and Ken Sochats, *Science Metadata Specifications Derived from the Functional Requirements: A Reference Model for Business Acceptable Communications*, available at <www.archimuse.com/papers/nhprc/BACartic.html>; note also the resource page on the Digital Curation Centre website: <www.dcc.ac.uk/resource/>.

7.6 Three terms are used in relation to the authentication of digital data, and a brief outline of each may be useful:

Authentication: this is the capacity to prove that the digital object is what it purports to be. The authenticity of a digital object is preserved by the use of techniques to prevent the data from being manipulated, altered or falsified deliberately or inadvertently. Such methods include providing audit trails of transmissions and maintaining records of encryption. A number of attributes, taken together, provide evidence of authenticity of the digital object: the mode, statue and form of its transmission, together with the way in which the data is preserved and how it is managed.

Integrity: this relates to how sound the data is, such as whether the data is

damaged in some way, and whether it is complete, in that it possesses all the necessary parts and links. Integrity is not an absolute condition, but is a state of relationships, and whether the burden of proof will be achieved in any individual case will depend on the strength of the relationships to the data.

Reliability: this is the capacity of a digital object to stand for the facts to which it purports to attest, which in turn is linked to ensuring sufficient procedural and technical attributes (including a combination of preventative measures, such as to prevent unauthorized amendments and changes, and verification measures to provide for a degree of assurance as to the identity of users and the provision of audit trails to the document when data is viewed and manipulated) are in place and working to provide for a degree of assurance that the digital object can be deemed to be reliable. In essence, reliability is associated with the degree of control exercised over the procedures that permit the data to be created. It is not absolute.

Challenges to the authenticity of electronic evidence

7.7 It is possible to challenge the authenticity of electronic evidence in a number of ways, although many reported cases appear to indicate that a lawyer will merely assert that the authenticity, reliability or accuracy of the evidence is not to be trusted, and the court will then have to determine a suitable response to the allegation raised,[1] or a lawyer may fail to raise any specific objections as to the accuracy of the evidence.[2] George L. Paul is of the opinion that the foundational requirement for authentication of electronic evidence has largely deteriorated into a 'trivial showing', and that without demonstrating that the information was created and stored within a reliable system, the continuity of custody necessary to show that an electronic document is authentic is lost.[3] The view that is taken is that the authenticity of electronic evidence is no trivial matter, and it goes beyond demonstrating its continuity of custody. A valid challenge to its authenticity may render the electronic evidence inadmissible, just as a robust defence of the authenticity of electronic evidence may preserve its admissibility.

1 For instance, in *Nobel Resources SA v Gross* [2009] EWHC 1435 (Comm), Mr Gross cast doubt over the reliability and (it seems) the authenticity of SMS messages, but the technical evidence demonstrated that it was not possible to alter an SMS message on a BlackBerry once it has been received or sent; note the discussion in relation to the print-outs of records of telephone calls made by a mobile telephone in the case of *State v Navjot Sandhu* (2005) 11 SCC 600, 148–152.

2 *Olympic Insurance Company v H. D. Harrison, Inc.*, 418 F.2d 669 (5th Cir. 1969).

3 George L Paul, *Foundations of Digital Evidence* (American Bar Association 2008), 48; George L Paul, 'Systems of Evidence in the Age of Complexity' (2014) 12 Ave Maria L Rev 173.

Types of challenges

7.8 Challenges to the authenticity of electronic evidence can include:

1. Claiming that the records were altered, manipulated or damaged between the time they were created and the time they appeared in court as evidence.[1]

2. Questioning the reliability of the program that generated the record.

3. Disputing the identity of the author of the electronic evidence: for instance, the person ostensibly responsible for writing a letter in the form of a word processing file, SMS or email may dispute he wrote the text, or sufficient evidence has not been adduced to demonstrate the nexus between the evidence and the person responsible for writing the communication.

4. Questioning the reliability of the evidence from a social networking website.

5. Even if it might be agreed that an act was carried out and recorded in an electronic message, failing to prove the message was directed to a particular person, especially where others might have access to the device (such as a mobile telephone) that produced the message.

6. Questioning whether the person alleged to have used his PIN, password or clicked the 'I accept' icon was the person who actually carried out the action.

1 Seng, 'Computer output as evidence', 163–9.

7.9 However, it is questionable whether a lawyer challenging electronic evidence can ever raise sufficient doubt about the authenticity of digital data because of the complexity of the systems and the difficulty of obtaining evidence from the various owners of the different part of any given system.[1] But this is not always the case, as shown in *Koosharem Corporation v SPEC Personnel, LLC*.[2] In this case, Koosharem alleged that the absence of authenticity of the digital documents tendered by SPEC justified the services of a digital evidence professional to interrogate SPEC's computers more thoroughly. It was alleged that the emails tendered were not accurate, because the date and time stamp on each email appeared to have been modified to reflect the dates the emails were compiled, rather than the dates they were sent, and these irregularities in the emails that were produced called into question the authenticity of the documents. In this instance, Catoe MJ determined that it was necessary to conduct a forensic analysis, and set out the procedure that the parties were required to follow in relation to the technical aspects of the analysis:

(1) Defendants will make available for forensic analysis and data recovery to be conducted by an expert forensics firm ('Expert') any business computers and/or any personal computers used to conduct business, correspond in any way regarding business, Spec and/or its current or employees, and/or plaintiffs and/or their current or former employees [individuals identified]

(2) The time frame for the forensic analysis and data recovery will encompass the period [starting period], to present.

(3) The parties will jointly agree within five (5) calendar days after entry of this order on an Expert that will be used to conduct the data recovery and forensic analysis.

(4) Defendants will produce to the Expert within ten (10) calendar days after entry of this order the computers identified in paragraph 1.

(5) The Expert will recover only the documents and email account or accounts used by individuals identified in paragraph 1 (or those accounts and documents accessed remotely using another computer).

(6) The Expert also will conduct a search or run other appropriate programs to determine whether any emails or documents have been deleted, destroyed, altered, or otherwise compromised since [date], and whether any programs have been installed that would alter, destroy, erase, modify, or otherwise compromise any portion of each computer or its contents as of [date]. The Expert also will be permitted to conduct such search efforts as are necessary to form an opinion as to whether any procedures were put into place to preserve emails and documents as of [date].

(7) The recovery of emails will include all emails in any form whatsoever including, but not limited to, deleted emails, forwarded emails, copied ('cc') and blind-copied ('bcc') emails and draft emails. The recovery of documents will

include all documents including drafts, multiple versions, and final versions.

(8) The Expert will securely maintain the original data recovered in order to establish a chain of custody.[3]

(9) The Expert will produce a copy of the recovered data to defendants' local counsel of record [as identified].[4]

1 Luciana Duranti and Corinne Rogers make this point, with further citations, in 'Trust in digital records: An increasingly cloudy legal area' (2012) 28 Computer Law and Security Review 522, 537.
2 2008 WL 4458864 (D.S.C. Sept. 29, 2008).
3 It is important to establish the continuity of evidence (also called the chain of custody) for data taken from devices, while it may not be held inadmissible, the court may consider it circumstantial and will need to rule on its admissibility: *R v Avanes* 2015 ONCJ 606.
4 2008 WL 4458864 (D.S.C. Sept. 29, 2008), pp 2-3.

7.10 It will be rare for a judge to set out in detail the nature of the action a party ought to undertake in civil proceedings to establish the authenticity of electronic documents. However, where a judge may consider it necessary to issue such an order, the instructions above as set out by Catoe MJ are a useful guide.

7.11 The way in which electronic evidence is adduced will affect the challenges as to its authenticity. Generally, evidence is adduced to assert or reinforce a positive position. For instance, it might provide reliable information,[1] act to confirm an alibi,[2] or where there is evidence from different devices and systems in combination (CCTV, automatic number plate recognition system and the use of mobile telephones[3] attributed to a particular person), act to corroborate and reinforce the evidence between the parties.[4] An example of the positive use of electronic evidence is the Application Transaction Counter on the chip on a debit card, which increases by one each time a transaction occurs, so that in the event of a disputed transaction, the counter on the card can be tested against the records maintained by the bank.[5] Another example of the positive use of electronic evidence can be found in the case of *City Park Co-operative Apartments Inc. v David Dubois.*[6] In this case, Spies J accepted that the code of an apartment entry-exit 'key' issued to the defendant contradicted the defendant's affidavit evidence that he had been denied access to his apartment. The management of the apartment was able to adduce in court evidence to show that this particular entry-exit 'key' was used 1,447 times in a six-month period, based on computer records of each entry or exit for the uniquely coded 'key' (which the judge questionably described as an 'electronic signature'). Nonetheless, the judge's meaning is clear: this was an example of electronic evidence demonstrating that the holder of a token had used the entry-exit 'key', thus going to show that his affidavit evidence was incompatible with the electronic evidence.

1 *A (Death of a Baby), Re* [2011] EWHC 2754 (Fam), per Jackson J at [168].
2 *Hallam, R. v* [2012] EWCA Crim 1158.
3 In *R. v Hamilton* 2011 ONCA 399 the Ontario Court of Appeal held that evidence regarding cell tower records was factual evidence, and not opinion evidence, and the court accepted evidence of three employees, rather than experts at [259]; see also *R. v Cyr 2012* ONCA 919; these decisions have been criticised by Chasse: Chasse, *Guilt by Mobile Phone Tracking Shouldn't Make 'Evidence to the Contrary' Impossible*.
4 *Fagan, R v* [2012] EWCA Crim 2248. Note the discussion of a case in Switzerland where the absence of evidence that a mobile telephone that was switched on at the relevant time was the topic of a paper in considering probability and graphical probability models: Alex Biedermann and Joëlle Vuille, 'Digital evidence, "absence" of data and ambiguous patterns of reasoning' (2016) 16 Digital Investigation S86–S95.
5 Jerzy Kosiński, 'A case of the customer attempting to claim their debit card was cloned' (2016) 13 Digital Evidence and Electronic Signature Law Review 167.

6 [2006] OJ No. 4428 (Sup. Ct.) (QL).

7.12 However, it is possible for digital data to prove a negative (or perhaps be adduced as evidence of an inconsistent positive), a point made by Professor Tapper.[1] An example is the case where a number of customers of a bank report unauthorized ATM withdrawals, which will cause the bank to investigate whether an employee was responsible for the thefts. This happened in the case of *United States of America v Bonallo*,[2] where computer records had demonstrated that cash withdrawals were made when the defendant Bonallo was in the building. It transpired that the employee who assumed Bonallo's duties after his employment was terminated discovered a 'fraud program' in Bonallo's computer program library. This program was used to provide him access to ATM computer files, and to alter transaction records, although it could have been used for legitimate purposes as well. This case not only illustrates the possibility of adducing evidence of an inconsistent positive, but also the care with which judges should approach assertions about 'reliable' computer systems and whether the business records exception ought to apply.

1 Colin Tapper, 'Evanescent evidence' (1993) 1 International Journal of Law and Information and Technology 35, 44–5; Beryl A Howell and Brian M Heberlig, 'The *Lamar Owens* Case: How Electronic Evidence Contributed to an Acquittal in an Explosive Rape Case' (2007) 24 The Computer & Internet Lawyer 1–4; *Alfano v LC Main, LLC*, 38 Misc.3d 1233(A) (2013) 969 N.Y.S.2d 801 (Table), 2013 WL 1111969 (N.Y.Sup.), 2013 N.Y. Slip Op. 50373(U) (a forensic computer examiner performed a forensic analysis of the metadata associated with plaintiffs' photographs, concluding that the photographs were taken 12 days after the accident); Kashmir Hill, 'Fitbit data just undermined a woman's rape claim' (*Fusion*, 29 June 2015) <http://fusion.net/story/158292/fitbit-data-just-undermined-a-womans-rape-claim/>.
2 858 F.2d 1427 (9th Cir. 1988).

Showing authenticity

7.13 In determining whether the party adducing the electronic evidence has discharged the burden of authenticating the evidence,[1] a range of factors may need to be taken into account, covering some or all of the technical attributes associated with the preservation of digital data.[2] The nature of the evidence required in civil matters will depend largely on the pleadings of the parties, which set out the issues in dispute, and the extent to which a party puts the other party to proof. Preparing and presenting evidence of the authenticity of digital data is a matter for the party who seeks to have the evidence admitted where the procedural rules require suitable evidential foundations to be met. This does not prevent both parties and the court from accepting the authenticity of the evidence without proof.[3] Alternatively, one party may put the authenticity or integrity of digital documents in issue. In such a case the party adducing the evidence will also need to meet the requirement to provide suitable evidential foundations. For instance, in civil proceedings in England and Wales, a party is deemed to admit the authenticity of a document disclosed under the provisions of CPR Pt 31 unless notice is served that the party wishes the document to be proved at trial, as provided for by CPR 32.19. Notwithstanding the provisions of the CPR, the authenticity of documents is not, generally, challenged at such an early stage in the proceedings.[4] This is because neither party may be aware of the dispute over the authenticity of a document until during the trial, when it may be first raised by a witness during oral testimony.

1 Statutory certification schemes are not an answer, given that once a witness (sometimes with questionable knowledge or competence to offer any opinion) testifies to the reliability of the device (whatever reliability means), it is then assumed, erroneously, that it is not necessary to provide any further evidence of authentication: Daniel Seng and Sriram S Chakravarthi, *Computer Output as Evidence: Consultation Paper* (Singapore Academy of Law 2003) 100–1.
2 For which see the discussion in Duranti and Rogers, 'Trust in digital records' 522.
3 For a number of early cases in the USA where digital images from satellites were accepted by agreement, see Harald Ginzky, 'Satellite Images as Evidence in Legal Proceedings relating to the Environment – A US Perspective' (2000) 25 Air & Space Law 114, 116.
4 Although see *Gallaher International Ltd v Tlais Enterprises Ltd (Rev 1)* [2008] EWHC 804 (Comm), where Gallaher gave notice that it challenged the authenticity of a large number of the documents disclosed by Tlais, and required Tlais to prove them, at [586]. Clarke J did not consider some documents proved or not proved, at [630], some were not proved [685], and some were not satisfactorily proved, at [862].

Guidelines and standards

7.14 When resolving the issue of the authenticity of electronic evidence, it is commonly thought that reference will be made to guidelines and standards, both national and international.[1] This was the opinion offered in the first edition of this text, and it is a view championed by Ken Chasse.[2] However, it is unlikely that much consideration will be given to any standards issued by national or international bodies, unless these are provided for in legislation or where a superior court directs judges to consider those standards that offer guidelines on system integrity and reliability, which is the case in Canada. For instance, the Alberta Evidence Act, RSA 2000, section 41.6 provides that:

> For the purpose of determining under any rule of law whether an electronic record is admissible, evidence may be presented in respect of any standard, procedure, usage or practice on how electronic records are to be recorded or stored, having regard to the type of business or endeavour that used, recorded or stored the electronic record and the nature and purpose of the electronic record.

1 See, for instance, *ISO/IEC 27037:2012 — Information technology — Security techniques — Guidelines for identification, collection, acquisition, and preservation of digital evidence*, (ISO/IEC 27037:2012(E), First edition, 2012-10-15); *Australia and New Zealand Guidelines for Digital Imaging Processes* (Electronic Evidence Specialist Advisory Group, 2013).
2 Ken Chasse, 'Electronic records as documentary evidence' (2007) 6 Canadian Journal of Law and Technology 141 and 'The admissibility of electronic business records' (2010) 8 Canadian Journal of Law and Technology 105 and 'Why a legal opinion is necessary for electronic records management systems' (2012) 9 Digital Evidence and Electronic Signature Law Review 17; 'legal reliability standards' were supported in the UK, for which see Stephen Castell, 'Evidence and authorisation: Is EDI "legally reliable"?' (1990) 6 Computer L & Secur Rep 4.

7.15 In *R v Oler*,[1] Lamoureux J of the Provincial Court of Alberta stated that the principles and procedures outlined by the Canadian General Standards protocol (specifically Standard CAN/CGSB-72.34-2005, Electronic Records as Documentary Evidence (the 'Standard')) 'is intended to enhance the potential for admissibility of electronic documents and reliance upon these documents as authentic for the purpose of business and Court proceedings',[2] and that 'records and documents including electronic images by or stored in a computer can stand in place of original paper source records or copies of paper source records'.[3] The judge cited paragraph 5.2.1 from the Standard:

> Those who wish to present an electronic record as evidence in legal proceedings shall be able to prove
>
> (a) authenticity of the record;
>
> (b) integrity of the [Records Management System (RMS)] that a record was recorded or stored in; and
>
> (c) that it is 'a record made in the usual and ordinary course of business' or that it is otherwise exempt from the legal rule barring hearsay evidence.
>
> …
>
> Therefore proof of the integrity of an electronic record is established by proof of the integrity of the RMS that recorded or stored it, which can be considered the 'system integrity test' of admissibility for electronic records. In the absence of evidence to the contrary, such system integrity can be proved by evidence that:
>
> a) the computer system was operating properly;[4] or if it was not, the fact of its not operating properly did not affect the integrity of the electronic record, and there are no other reasonable grounds to doubt the integrity of the electronic records system;
>
> b) the electronic record was recorded or stored by a party to the proceedings that is adverse in interest to the party seeking to introduce it as evidence; or
>
> c) the electronic record was recorded or stored 'in the usual and ordinary course of business by a person who is not a party to the proceedings and who did not record or store it under the control of the party seeking to introduce the record' as evidence;
>
> d) a printout of an electronic record that has been 'manifestly or consistently acted on, relied upon, or used as the record of the information recorded or stored on the printout is the record for the purposes of the best evidence rule,' i.e., the best evidence rule for paper documents applies since the printout is not being used to show the contents of a computer. In Alberta and Ontario, the evidence acts allow a party to support the integrity of the electronic record by showing that a reliable encryption system was used to create it.[5]

1 2014 ABPC 130.
2 2014 ABPC 130, [6].
3 2014 ABPC 130, [6].
4 Alternatively, where a system is operating consistently, or to put it another way, whether the system was working in accordance with an expectation, or the system returned generally verifiably correct results.
5 2014 ABPC 130, [7].

7.16 The judge went on to cite paragraph 5.5 of the Standard:

> The following can be used to prove an organization's usual and ordinary course of business, the integrity of its electronic records system and, therefore, the integrity of any record recorded or stored in that system:
>
> b) Contemporaneous recording: The electronic records were captured and recorded contemporaneously with, or within a reasonable time after, the events to which they relate (but contemporaneous recording within a particular data base is not required);
>
> c) Routine business data: The data within a record is of a type regularly supplied to the originating organization or created by it during its regular activities;
>
> d) Data entry: The data-base capture and entry procedures are part of the usual and ordinary course of business of the organization and are carried out in accordance with the procedures manual;

> e) Industry and national standards: The organization conforms to all appropriate standards for an RMS: inputting, importing and storing data, and preserving the reliability of data and of the RMS that stores and transmits the data;
>
> f) Business reliance: The organization, when making business decisions, relies upon the electronic records in its data bases;
>
> g) Software reliability: The software reliably processes the data;
>
> h) Recording of system changes: A record of system changes is kept; and,
>
> j) Security: Security features are used to guarantee the integrity of the RMS; at least, the following security features should be able to be proved:
>
> > 1) protection against unauthorized access to data and permanent records;
> >
> > 2) processing verification of data and information in records;
> >
> > 3) safeguarding of communications lines;
> >
> > 4) maintenance of backup copies of records to replace falsified, lost and destroyed permanent or temporary records
> >
> > 5) retention and disposition of electronic records in the compliance with legislated and internal retention periods and disposition requirements, and documenting such compliance and disposition schedules; and
> >
> > 6) establishment of a business continuity plan for electronic records and associated data, including off-site copies of essential files, operating and application software.[1]

1 2014 ABPC 130, [7].

7.17 Lamoureux J considered that the above factors set out in the standard could be:

> … proved by a single supervising officer of the organization who is accountable for the records system. An additional witness may be required for software unique to the system unless the supervisor can prove its history of reliability. If not, the programmer who wrote the software should be available to certify its reliability until the software does have a history of reliability. The programmer or developer shall obtain a security clearance from the organization. Therefore, in choosing suppliers and programmers, consideration should be given to their ability and experience to prove the reliability of their products. And in legal proceedings, the use of data from an electronic record should not violate any legal principles prohibiting the disclosure of privileged or confidential data.[1]

1 2014 ABPC 130, [7].

7.18 In essence, the judge in *R v Oler* considered that the 'reliability of electronic records for purposes of Court proceedings will be fundamentally linked to the Records Management System (RMS)'.[1] In this case, the electronic evidence took the form of the contents of some handwritten notes that were rekeyed into a Microsoft Word format, as well as some notes that were captured in the form of scanned Portable Document Format (PDF) documents. The judge considered evidence from several witnesses who were responsible for entering information into the RMS and for the security of the data. The judge agreed that it is 'critical to the integrity of the … system that the original document to be scanned into a PDF format be an exact reproduction of the original'.[2] Critically, information that was copied into a Word document from a handwritten note was not identical to the handwritten note. In conclusion, the judge observed that the Word document is a 'live document, which is capable of being changed from time to

time'.[3] Such documents could be admitted, but it was for the trial judge to determine what weight it was to be given, because it did not meet the standards for admission as an electronic document under CAN/CGSB 72.11-93 or CAN/CGSB 7234-2005.

1 2014 ABPC 130, [11].
2 2014 ABPC 130, [21].
3 2014 ABPC 130, [28].

7.19 The nature of the evidence available to a court to determine the authenticity of digital data will differ from case to case, as indicated by Lord Griffith in *R v Shephard*:

> Computers vary immensely in their complexity and in the operations they perform. The nature of the evidence to discharge the burden of showing that there has been no improper use of the computer and that it was operating properly will inevitably vary from case to case. I suspect that it will very rarely be necessary to call an expert and that in the vast majority of cases it will be possible to discharge the burden by calling a witness who is familiar with the operation of the computer in the sense of knowing what the computer is required to do and who can say that it is doing it properly.[1]

1 [1993] AC 380, 387; [1993] 1 All ER 225 (spelt 'Shepherd' in All ER); [1993] Crim LR 295, HL. See *Connolly v Lancashire County Council* [1994] RTR 79, QBD, where the prosecution elected to produce evidence that a weighbridge was working properly, but failed to demonstrate the computer was functioning properly at the material time.

7.20 In this passage, Lord Griffiths was referring to the requirement to comply with s 69 of the Police and Criminal Evidence Act 1984. Under s 69, a statement in a document produced by a computer was not admissible as evidence in criminal legal proceedings unless it could be shown that there were no reasonable grounds for believing that the statement was inaccurate because of improper use of the computer, and the computer was operating properly (whatever 'operating properly' means) at all times or, if not, that any respect in which it was not operating properly or was out of operation did not affect the production of the document or the accuracy of its contents. Before a judge could decide whether computer print-outs were admissible as evidence, it was necessary to call appropriate authoritative evidence to describe the function and operation of the computer. This normally consisted of a statement of evidence as to how the print-out was obtained, together with a certificate signed by a person occupying a responsible position in relation to the operation of the computer to the effect that the computer system was operating correctly at the time the evidence was obtained. Although s 69 was repealed by s 60 of the Youth Justice and Criminal Evidence Act 1999, nevertheless the comment is a useful reminder that when there is a requirement to prove the reliability of a digital document, it will be necessary to ensure the relevant witnesses are qualified to offer the requisite evidence. This must be right.

7.21 According to the author of 'Admissibility of electronically filed federal records as evidence',[1] in 1990, cross-examination in relation to the integrity of computer stored or generated files included questioning the source of the input data or information and the process for transcribing it to machine readable form, the computer programs that create, edit and update the files, the computer programs that produce the output or stored files, and the reliability of the hardware and vendor-supplied 'off-the-shelf' software that systematically manages the internal processes of the computer. In this

respect, the lawyer whose duty it is to test the evidence is not interested in the gradual build-up of the various layers of technical and organizational characteristics that form the basis for the authenticity of data in digital form. She is interested in exposing weaknesses in the evidence, and if it can be demonstrated that a sufficient number of weaknesses exist, the totality of the cross-examination may mean that the party submitting the document has failed to discharge the evidential burden of convincing the adjudicator to accept the evidence.[2] Procedures, process and technical measures such as audit logs, system security and the use of digital signatures are all highly relevant in providing for the authenticity of digital data, as are the methods by which people are required to interact with computers and computer systems.

1 IV, Conclusion, (US Department of Justice, October 1990), available at <www.lectlaw.com/files/crf03.htm>.

2 For a discussion relating to email, see Chris Reed, 'Authenticating electronic mail messages – some evidential problems' (1989) 52 Modern Law Review 649; Mark D Robbins, 'Evidence at the electronic frontier: Introducing e-mail at trial in commercial litigation' (2003) 29 Rutgers Computer and Technology Law Journal 219.

7.22 However, manipulation of computers and computer systems can and do occur notwithstanding the existence of documented procedures and processes as guidelines and standards. By way of example, in March 2012 the US federal government and state attorneys general filed consent judgments in the US District Court in the District of Columbia with the Bank of America Corporation, J.P. Morgan Chase & Co, Wells Fargo & Company, Citigroup Inc and Ally Financial Inc, to resolve violations of state and federal law in relation to foreclosure documents. It transpired even though documented processes exist, a range of incorrect practices had occurred in breach of these processes, including notaries signing and stamping documents that had not been signed, employees signing affidavits that included information that was not within their personal knowledge, employees signing documents under a Certificate of Incumbency as if they held office under the relevant legal entity they were signing for, and employees deliberately creating false computer customer files to hide the truth and altering data on computer systems.[1] The development, provision of standards and guidelines and their implementation are merely one part of the whole analysis. The gap between what is stated in the standard or guideline, and what actually occurs in reality, will be a central focus of cross-examination in a court.[2]

1 For more information, including copies of the relevant documents, see <www.nationalmortgagesettlement.com>; <www.justice.gov/opa/pr/25-billion-mortgage-servicing-agreement-filed-federal-court>.

2 For instance, the issuing of certificates to accompany digital signatures, which is a foundation of the security of the Internet, is bound by strict legal requirements in most legislation, yet DigiNotar B.V., a Dutch certificate authority, was put into liquidation by the Dutch government for issuing false certificates, for which see Stephen Mason, *Electronic Signatures in Law* (4th edn, University of London 2016), 15.15.

Judicial approaches to authentication

7.23 The authenticity of digital data in legal proceedings is considered on a case by case basis. There are no extensive guidelines about the attributes or characteristics of digital data, and some commentators have provided guidance through the application of relevant case law in relation to different types of digital data, such as emails, web

sites, instant messages and text messages and photography.[1] The Federal Judicial Centre in the US amended the *Manual for Complex Litigation* to accommodate the nature of evidence in digital form:

> *Use at trial.* In general, the Federal Rules of Evidence apply to computerized data as they do to other types of evidence. Computerized data, however, raise unique issues concerning accuracy and authenticity. Accuracy may be impaired by incomplete data entry, mistakes in output instructions, programming errors, damage and contamination of storage media, power outages, and equipment malfunctions. The integrity of data may also be compromised in the course of discovery by improper search and retrieval techniques, data conversion, or mishandling. The proponent of computerized evidence has the burden of laying a proper foundation by establishing its accuracy.
>
> The judge should therefore consider the accuracy and reliability of computerized evidence, including any necessary discovery during pretrial proceedings, so that challenges to the evidence are not made for the first time at trial. When the data are voluminous, verification and correction of all items may not be feasible. In such cases, verification may be made of a sample of the data. Instead of correcting the errors detected in the sample—which might lead to the erroneous representation that the compilation is free from error—evidence may be offered (or stipulations made), by way of extrapolation from the sample, of the effect of the observed errors on the entire compilation.
>
> Alternatively, it may be feasible to use statistical methods to determine the probability and range of error.[2]

1 Steven Goode, 'The Admissibility of Electronic Evidence' (2009) 29 Review of Litigation 1; Breanne M Democko, 'Social Media and the Rules on Authentication' (2012) 43 U Tol L Rev 367; Kenneth N Rashbaum, Matthew F Knouff and Dominique Murray, 'Admissibility of Non-U.S. Electronic Evidence' (2012) 18 Rich J L & Tech 9; Paul W Grimm, Lisa Yurwit Bergstrom and Melissa M O'Toole-Loureiro, 'Authentication of Social Media Evidence' (2013) 36 American Journal of Trial Advocacy 433.

2 *Manual for Complex Litigation* Fourth (2004), Federal Judicial Center, 11.446, 82. Original footnote omitted.

7.24 Judges have to make judgments about the qualifications of the witnesses who appear before them, and interpret the nature of digital data in accordance with the evidence presented. Two cases from the US and one from England and Wales serve to illustrate this point. In *State of New Jersey v Swed*,[1] the defendant was convicted of obtaining electricity without payment, and part of the evidence comprised computer print-outs identifying the defendant as a customer, with a registered address. The defendant contended that there was insufficient foundation for the admission of the print-outs. In reaching its decision, the Appellate Division of the Superior Court of New Jersey applied the six foundational requirements set out in *Monarch Federal Savings & Loan Association v Gesner*:[2] personal knowledge on the part of the witness as to the act or event recorded was not necessary; the person called as a witness should be able to testify as to the type of computer used, the permanent nature of the record storage, and how daily transactions were customarily recorded; the computer records were made in the ordinary course of business; the entries were made within a reasonable time after the transaction occurred; proof of the validity of the source of the information from which the entry was made; and the validity of the method used in obtaining the computer print-out. In each of these instances, the prosecution provided suitable evidence. Gruccio JAD went on to comment:

> With the advent of computers has come an implicit trust in their dependability, owing primarily to the results they achieve. The mechanical (or electronic) explanation of computer workings would likely have been beyond the grasp of most jury members and would not have proved helpful in establishing the reliability of the records … An explanation of the internal workings of a massive computer system belies common sense and judicial efficiency[1]. Computer usage permeates every strata of society and is customary in modern life.[3]

1 604 A.2d 978 (N.J.Super.A.D. 1992).
2 156 N.J.Super. 107, 383 A.2d 475 (Ch.Div. 1977).
3 At 982–3.

7.25 In footnote [1], the judge commented: 'Some thought should be given, however, to the time period in which the *Monarch* opinion was penned, and the progress that has been made in computer technology over the past 15 years. The requirement that a modern, foundation witness testify to a computer's mechanical workings is ineffectual in proving its reliability.' Although the observations by the judge in relation to the perceived overall reliability of computers were not accurate, the comment by Gruccio JAD that it was impossible to provide meaningful evidence of the workings of a computer for the purposes of a foundation must be correct – the test of authenticity cannot be founded on how the software code interacts, but whether the evidence tendered satisfies the tests set out below.

7.26 In the criminal case of *R v Cochrane*,[1] McCowan LJ, Waterhouse and Brooke JJ set out the following guidance in relation to electronic evidence from mainframe computers:

> … it was necessary that appropriate authoritative evidence should be called to describe the function and operation of the mainframe computer, including the extent to which it brought to bear information stored within it in order to validate a transaction and to enable an appropriate record to be made on the till roll. None of those matters were covered by any of the witnesses, and the judge had had to grapple with inadequate, and possibly, incorrect information. …. The Crown had failed to adduce adequate evidence to enable the court to properly rule that the till rolls were admissible evidence; and in the absence of the till rolls the Crown's case could not be proved.

1 [1993] Crim LR 48 Court of Appeal, Criminal Division.

7.27 In the context of the US, George L. Paul has indicated that '… the Federal Rules of Evidence do not contain a rule requiring informational records or other objects to be authentic. The requirement appears to be assumed …',[1] and he indicates that authenticity is a *prerequisite,* because evidence must also be relevant. This observation must be considered to be accurate for most jurisdictions. In this part of the chapter, consideration is given to a number of jurisdictions and how judges have approached the authentication of digital data, and to illustrate that comprehensive tests to demonstrate the authenticity of digital data are not necessary for every conceivable set of facts – an observation made by Erdmann J in *United States v Lubich* before the U.S. Court of Appeals for the Armed Forces, in which he said 'There are numerous scenarios in which this issue will arise and we see no benefit in attempting to craft a "standard" test to analyze all computer data situations'.[2]

1 Paul, *Foundations of Digital Evidence*, 39; Paul, 'Systems of Evidence in the Age of Complexity'; see an earlier comment: Rudolph J Peritz, 'Computer Data and Reliability: A Call for Authentication

of Business Records Under the Federal Rules of Evidence' (1986) 80 Northwestern University Law Review 965.

2 72 M.J. 170 (2013), 175 – the attorney for the appellant argued that the prosecution had failed to provide for the continuity of the evidence.

7.28 The US case of *In re Vee Vinhnee, debtor, American Express Travel Related Services Company, Inc. v Vee Vinhnee*[1] excited a renewed interest in the foundational requirements of electronic evidence. The case dealt with a failure to introduce a sufficient evidentiary foundation for the introduction of business records in digital form. In this case, American Express claimed Vinhnee failed to pay credit card debts, and took action to recover the money. After a trial before the Bankruptcy Court for the Central District of California that occurred in the absence of the defendant, the trial judge determined that American Express failed to authenticate certain records in digital form. American Express appealed the verdict, and the decision of the trial judge was affirmed. In respect of the issues in this particular trial, Klein J pointed out that:

> the focus is not on the circumstances of the creation of the record, but rather on the circumstances of the preservation of the record during the time it is in the file so as to assure that the document being proffered is the same as the document that originally was created.[2]

1 336 B.R. 437 (9th Cir. BAP 2005).

2 336 B.R. 437 (9th Cir. BAP 2005), 444 [14].

7.29 The judge made the pertinent point that: 'Ultimately, however, it all boils down to the same question of assurance that the record is what it purports to be.'[1] The judge continued to explain the issues involved in this process:

> The logical questions extend beyond the identification of the particular computer equipment and programs used. The entity's policies and procedures for the use of the equipment, database, and programs are important. How access to the pertinent database is controlled and, separately, how access to the specific program is controlled are important questions. How changes in the database are logged or recorded, as well as the structure and implementation of back-up systems and audit procedures for assuring the continuing integrity of the database, are pertinent to the question of whether records have been changed since their creation.
>
> There is little mystery to this. All of these questions are recognizable as analogous to similar questions that may be asked regarding paper files: policy and procedure for access and for making corrections, as well as the risk of tampering. But the increasing complexity of ever-developing computer technology necessitates more precise focus.[2]

1 At 445 [15].

2 336 B.R. 437 (9th Cir. BAP 2005), 444 [16].

7.30 Klein J reached the conclusion that early attempts at establishing a foundation for electronic evidence were too cursory, while also accepting that judicial notice is commonly taken of the validity of the theory underlying the use of computers and the validity of the data generated generally. The judge then set out the tests described by Professor Imwinkelried when considering electronic records as a form of scientific evidence:[1]

1. The business uses a computer.
2. The computer is reliable.
3. The business has developed a procedure for inserting data into the computer.
4. The procedure has built-in safeguards to ensure accuracy and identify errors.
5. The business keeps the computer in a good state of repair.
6. The witness had the computer readout certain data.
7. The witness used the proper procedures to obtain the readout.
8. The computer was in working order at the time the witness obtained the readout.
9. The witness recognizes the exhibit as the readout.
10. The witness explains how he or she recognizes the readout.
11. If the readout contains strange symbols or terms the witness explains the meaning of the symbols or terms for the trier of fact.

1 Edward J Imwinkelried, *Evidentiary Foundations* (9th edn, LexisNexis 2015) 4.03[2].

7.31 The steps outlined by Professor Imwinkelried are helpful, but item 1 is hardly a ground for admitting electronic evidence, in that software in computers and computer-like devices are put on the market when a manufacturer is satisfied that such devices will sell, not that they are reliable or can be trusted to be accurate. Item 2 is almost impossible to demonstrate,[1] and item 5 is prone to being undermined by the failure of an organization to consider such issues when operating its computers, although it is debatable whether the concepts of a computer being reliable or in a good state of repair are helpful (or relevant) in understanding whether a computer was working properly – and the term 'working properly' is also to be questioned. Chasse refers to this list as superficial,[2] and argues that it is less demanding than the test set out in the Canadian case of *R v McMullen*,[3] in which Morden JA set out what can be described as the forerunner to the 'system integrity test':

> The nature and quality of the evidence put before the Court has to reflect the facts of the complete record keeping process – in the case of computer records, the procedures and processes relating to the input of entries, storage of information, and its retrieval and presentation: …. If such evidence be beyond the ken of the manager, accountant or the officer responsible for the records … then a failure to comply with s. 29(2) must result and the print-out evidence would be inadmissible.[4]

1 Please see Chapter 6: The presumption that computers are reliable.
2 Chasse, 'The Admissibility of Electronic Business Records'; Chasse also refers to the nine points of proof specified in the National Standard of Canada, *Electronic Records as Documentary Evidence* CAN/CGSB-72.34-2005, section 5.5.
3 [1979] OJ No. 4300, (1979), 25 OR (2d) 301, 47 CCC (2d) 499 at 506, 100 DLR (3d) 671 (Ont. C.A.).
4 100 DLR (3d) 671 (Ont. C.A.), 678–679.

7.32 Chasse suggests[1] that the authentication rule at times appears inadequate, because it cannot be established that an electronic record is the same as its first instantiation simply by looking at the record itself.[2] Luciana Duranti and colleagues posit that the Canadian Uniform Electronic Evidence Act renders it necessary to refer to an unbroken line of traces left by all those who interacted with the record or to the legitimate custody of a professional who can account for them,[3] suggesting that the weight is on the integrity of the system, rather than the record. According to them, the

authentication rule, at times, appears inadequate, because 'originality cannot easily be established.'

1 Chasse, 'The Admissibility of Electronic Business Records', 111.
2 Indeed, this concept was reiterated in *R v Nardi* 2012 BCPC 0318 where the court held that in order to support a finding that electronic records are authentic and the 'best evidence' of the information proffered, the party seeking to admit the evidence 'cannot simply look to the documents themselves'. Further, the court said this is especially so when considering information generated from a novel 'system'.
3 Luciana Duranti, Corinne Rogers and Anthony Sheppard, 'Electronic Records and the Law of Evidence in Canada: The Uniform Electronic Evidence Act Twelve Years Later' (2010) 70 Archivaria 95, 98; see also Heather MacNeil, 'Providing Grounds for Trust: Developing Conceptual Requirements for the Long-term Preservation of Electronic Records' (2000) 50 Archivaria 52; Luciana Duranti and Kenneth Thibodeau, 'The Concept of Record in Interactive, Experiential and Dynamic Environments: the View of InterPARES' (2006) 6 Archival Science 13; Luciana Duranti, 'From Digital Diplomatics to Digital Records Forensics' (2009) 68 Archivaria 39.

7.33 In addition, even moderately-sized organizations have systems as collections of computers, rather than single computers, in place, and the reliability and repair (if these terms, inaccurate as they are, are to be used) of such systems will differ markedly. To consider just one example, if the security patches are not kept up to date, the reliability of the system can be seriously undermined.[1] In *In re Vee Vinhnee*, Klein J amplified the fourth step in Professor Imwinkelried's test to include additional highly relevant factors as follows:

> The 'built-in safeguards to ensure accuracy and identify errors' in the fourth step subsume details regarding computer policy and system control procedures, including control of access to the database, control of access to the program, recording and logging of changes, back-up practices, and audit procedures to assure the continuing integrity of the records.[2]

1 The range of technical issues that can adversely affect any computer or network of computers: Daniel Bilar, 'Known knowns, known unknowns and unknown unknowns: anti-virus issues, malicious software and Internet attacks for non-technical audiences' (2009) 6 Digital Evidence and Electronic Signature Law Review 123.
2 336 B.R. 437 (9th Cir. BAP 2005), 444 [16].

7.34 The judge then proceeded to evaluate the exhibits submitted by American Express using the tests set out by Professor Imwinkelried. The evidence of the custodian of the records at American Express was far too vague to be accepted. The following problems were identified: generally, the evidence was unclear and unpersuasive; the custodian did not have the requisite knowledge to provide the evidence; the person providing evidence on behalf of American Express merely asserted that he was an employee of American Express and was personally familiar with the systems, without informing the court of his job title or of his relevant experience and training that would provide an element of authority to his evidence; American Express failed to provide information about its computer policy and system control procedures, control of access to the relevant databases and to the applicable programs, how changes to the data were recorded or logged, what back-up practices were in place, and whether there were any audit procedures used to provide assurance of the continuing integrity of the records.

7.35 The careful review of Klein J in *In re Vee Vinhnee* demonstrates that the nature of the testimony to support the authentication of electronic evidence will differ, according

to the nature of the evidence, such as whether the source is analogue or digital, and the nature of the electronic evidence that is to be authenticated: evidence of the use of an ATM or cash card, pages from Internet websites, email correspondence, chat rooms,[1] instant messaging sessions,[2] to mention a few of the more obvious forms of evidence in digital form. Although much of the evidence is admitted based on the testimony of one or both of the parties, nevertheless in some circumstances it is necessary to provide appropriate evidence of the methods used to record and store data, such as instant message communications, as in *U.S. v Jackson*.[3] In this case, an agent of the Postal Investigation Service, David Margritz, acting under cover, posed as a fourteen-year-old girl and entered into 'chat' conversations with the defendant. It was agreed by both parties that there were no original transcripts of the conversations, no original print-outs, or copies on floppy discs, hard drives or disc drives that recorded the conversations. None of the conversations were saved. Apparently Mr Margritz wiped his computer clean during a routine upgrade a year or two after the investigation. As a result, the government sought to introduce the notes taken by Mr Margritz from the online chats and saved into a Microsoft Word document. Bataillon CJ set out the nature of the evidence:

> He further testified that, at the end of each chat session, he saved the conversations between k8tee4fun and gnesta18 by clicking and dragging to highlight the complete conversation from start to finish. … He then copied and pasted the entire selection into a word processing document in Microsoft Word. … He testified that he saved each conversation chronologically in an ongoing log. … He further testified that immediately after he copied and pasted the conversations into Word, he made another copy for himself and added certain notes and edits to that copy. … He acknowledged that it was possible to leave out words if they were not properly highlighted and dragged, but stated that there was no human error in this case because he took 'great pains' to look back at the screen and make sure he captured everything accurately before closing the chat window. … He further testified that he never modified the document in any way. … He testified that he never relied on the archives of Yahoo, apparently because it was unavailable or he had been told it was not reliable.[4]

1 Examples from the US include *United States v Simpson*, 152 F.3d 1241 (10th Cir. 1998) (a combination of identifying information that the user gave in the chat and corroborating evidence found in the defendants home near his computer was sufficient to authenticate the chat log); *United States v Tank*, 200 F.3d 627 (9th Cir. 2000) (evidence of another chat room user that he recorded the chats and printed them out – the print-out appeared to accurately represent the chats and was sufficient to establish prima facie authenticity); *United States v Gagliardi*, 506 F.3d 140 (2nd Cir, 2007) (the informant and the agent testified that the exhibits were accurate records of the chat conversations); *United States v Barlow*, 568 F.3d 215 (5th Cir. May 6, 2009) (the appeal court decided that the testimony of the witness was sufficient to authenticate the chat log presented at trial).

2 For instance, in *Adams v Disbennett*, 2008 WL 4615623 (Ohio App. 3 Dist., Oct 20, 2008), the trial judge permitted Mr Adams to authenticate the record of the various chats through his own testimony.

3 488 F.Supp.2d 866, 73 Fed. R. Evid. Serv. 959.

4 488 F.Supp.2d 866, 869.

7.36 A digital evidence professional gave evidence that it was more appropriate to take a forensic copy of the hard drive to confirm the communications that were recorded, and if that was not possible, then there were other ways to accurately save computer chats, such as log files are saved to the hard drive, the use of the 'ypager' log found in Yahoo, third-party software programs available that would accurately save online chats, the basic 'print screen' and 'file-print' options that would have captured

the entire chat. The specialist proffered the opinion that the method employed by Mr Margritz was the least effective way to record the chat log. As a result of the observations by the digital evidence professional and the fact that there were examples of missing data, timing sequences that did not make sense, and other editorial information, the judge concluded that the document did not accurately represent the entire conversations that took place between the defendant and Mr Margritz. For this reason, the document was held to be inadmissible.

7.37 In 1969, Gillespie J set out criteria for authenticating evidence in the form of testimony recorded on magnetic tapes in the case of *King v State of Mississippi for Use and Benefit of Murdock Acceptance Corporation*,[1] before the Supreme Court of Mississippi. Print-out sheets of business records stored on magnetic tapes were admissible in evidence if it is shown:

> (1) that the electronic computing equipment is recognized as standard equipment, (2) the entries are made in the regular course of a business at or reasonably near the time of the happening of the event recorded, (3) the foundation testimony satisfies the court that the sources of information, method and time of preparation were such as to indicate its trustworthiness and justify its admission.[2]

1 Miss., 222 So.2d 393.

2 Miss., 222 So.2d 393, [9]. See *United States of America v Weatherspoon*, 581 F.2d 595 (7th Cir, 1978) and *Rosenberg v Collins*, 624 F.2d 659 (1980), where sufficient testimonial evidence was adduced to lay the foundations for the admission of computer print-outs.

7.38 Gillespie J continued to indicate that he did not consider that the evidence produced from a computer is always correct: 'We are not to be understood as indicating that computer evidence is infallible. Its probative value is the same as conventional books, and it is subject to refutation to the same extent.'[1] In 1972, Abrahamson, Seidenfeld and Guild JJ of the Appellant Court of Illinois, Second District, in *People of the State of Illinois v Gauer*[2] reached a similar conclusion.[3] In the Australian criminal case of *R v Chen*,[4] case law from England & Wales, the United States and Australia was considered in relation to the tape recordings of conversations. The issue was whether any conditions should be met in respect of conversations that were recorded on tape before being admitted. The members of the Court of Criminal Appeal of the Supreme Court of Victoria said that it depended on the circumstances of each case:

> The test is whether there is sufficient material before the court to allow the tribunal of fact acting reasonably to conclude that the recorded sounds reproduce those originally made by the persons identified by the evidence. In other words, there must be evidence, which the tribunal of fact is entitled to accept, that the recording is of a conversation which occurred and which would be admissible if proved by oral testimony. In our opinion, admissibility does not depend on the party tendering the tapes having removed absolutely any chance that they are inaccurate.[5]

1 Miss., 222 So.2d 393, [10].

2 7 Ill.App.3d 512, 288 N.E.2d 24.

3 The three tests were applied from the 1986 case of *Victory Memorial Hospital v Rice*, 493 N.E.2d 117 (Ill.App. 2 Dist. 1986) – Lindberg J indicated at 121 in *People of the State of Illinois v Gauer*, 7 Ill. App.3d 512, 288 N.E.2d 24 that 'this court stated that considering the general use of electronic computing and recording equipment in the business world and the business world's reliance on the equipment, the scientific reliability of such machines can scarcely be questioned.' The court in *Gauer* did not make such a statement. Abrahamson J quoted these words from *Jones on Evidence*, 5th edition, s 609.

4 [1993] 2 VR 139.
5 [1993] 2 VR 139, 150.

7.39 Although it will not be relevant or necessary to provide such an in-depth analysis of electronic evidence in every case brought before a court, nevertheless, as the review of case law from the various jurisdictions confirms, the comments made by Klein J in *In re Vee Vinhnee* are highly pertinent as they illustrate the nature of the authentication evidence that should be gathered, if it is necessary to adduce such evidence.

Self-authentication

7.40 There is some controversy in Australia regarding whether a document is capable of authenticating itself. In the Australian case of *National Australia Bank Ltd v Rusu*,[1] the National Australia Bank (NAB) claimed that Ms Rusu stole a large sum of money from it with the assistance of a Mr Mato. To assist their case, NAB alleged that after stealing the money, Rusu and Mato had available to them a significant amount of funds compared with their previous resources. NAB sought to recover the money allegedly stolen, and to assert charges and obtain tracing orders over their assets. NAB tendered two pages of what appeared to be a transaction history inquiry in relation to an account identified by a number. However, nothing on the face of these pages identified the bank or the customer. There was evidence that these pages were in a bundle of documents produced by Advance Bank in response to a subpoena that specified bank records for a different period, and to the effect that Advance Bank's customer was Mr Mato, whose full name was Peter Francis Mato. A solicitor for NAB made an affidavit attaching a schedule of payments, alleging that Mr Mato had paid a substantial sum of money into the Advance Bank account on the day after the alleged theft.

1 (1999) 47 NSWLR 309.

7.41 Bryson J rejected the tender of the two pages that purported to be part of the bank statement. The judge carefully distinguished between the authentication of documents, their relevance as evidence, the procedure for proving the contents of documents, and the admissibility of representations in documents as business records notwithstanding the hearsay rule. Bryson J held it was necessary to establish by evidence, other than the documents themselves, that the pages were a bank statement, which comprised a statement of Advance Bank and that the account to which they referred was an account of Mr Mato. The judge rejected the idea that under the Evidence Act 1995 (NSW), the authenticity of a document tendered in evidence could be determined simply on the basis of the form and content of the document itself, or alternatively, on the basis of that document with information about the source of the document, which in this case, showed that it was produced on subpoena and the identity of the person who produced it.

7.42 *NAB v Rusu* was cited with approval in subsequent cases.[1] However, Bryson J's reasoning in *NAB v Rusu* was criticised by Stephen Odgers SC[2] in the fifth edition of *Uniform Evidence Law*. Odgers, after quoting Bryson J's observation that the question of authenticity is not a question as to the relevance of documents within s 58(1), inferred that on Bryson J's approach, the court may not draw reasonable inferences from a document as to its authenticity. Odgers suggested that such a view is inconsistent with the intention behind s 58(1) and its legislative history. In *Lee v Minister for Immigration*

& Multicultural & Indigenous Affairs,[3] Madgwick J took up Odgers' criticism. The judge described the decision in *NAB v Rusu* as a 'controversial NSW authority', concluding that Bryson J may have meant

> … no more than that there may be cases in which, as a matter of fact, no inference as to authenticity of a document may be properly drawn from the document itself. If he meant to say more than that, it is by no means clear to me that the way is open for a court to read some unexpressed limitation into a grant of power to courts: such grants are generally very liberally construed … Such an approach may be particularly apt where, as here, the provision aims at putting another nail in the coffin of unmeritorious technicality in litigation and s 135 provides ample safeguards against possible abuse of the section.[4]

1 In *Daw v Toyworld (NSW) Pty Ltd* (2001) 21 NSWCCR 389, the appellant brought proceedings for damages for a workplace injury. One of the grounds of appeal was that the trial judge had erred in placing reliance on a set of clinical notes of unknown origin. Heydon JA (with whom Priestley and Sheller JJA agreed) rejected this ground because it had not been shown that the trial judge placed reliance on this material, and no objection to its admissibility had been taken at the trial. The judge added '... if the document was of unknown origin, it could have been objected to as unauthenticated and irrelevant. The Evidence Act 1995 does not permit documents to authenticate themselves save in limited circumstances [citing *NAB v Rusu*]). See also *Kingham v Sutton (No 3)* [2001] FCA 1117 (15 August 2001) [127] (Goldberg J) and *Citibank Ltd v Chiu Wah Liu* [2003] NSWSC 69 [5] (Hamilton J). In *Crime Commission (NSW) v Trinh* [2003] NSWSC 811 (5 September 2003) Hidden J at [14] drew attention to the distinction between authenticity of records and accuracy of records. The judge distinguished *NAB v Rusu*, while not disagreeing with it, on the ground that the argument before him, relating to some casino records, was concerned with their accuracy rather than authenticity.

2 Stephen Odgers, *Uniform Evidence Law* (Lawbook Co 2002).

3 [2002] FCAFC 305 (4 October 2002).

4 [2002] FCAFC 305 (4 October 2002), [85]. See also *Albrighton v Royal Price Alfred Hospital* (1980) 2 NSWLR 542.

7.43 In *ASIC v Rich*,[1] documents from a file server had been tendered as evidence, and their provenance was questioned. After consideration of the authorities and of the evidence before him, Austin J concluded that there were sufficient grounds to authenticate each category of documents, which originated on file servers at One. Tel. In arguing against authentication, the defendants made two general submissions about authentication that the judge considered. First, the defendants submitted that the fundamental problem with all categories of documents was that ASIC had brought forward no one who was involved in the creation or keeping of the documents who could verify that they were final and operative documents as they existed at any particular point of time, as opposed to merely being some drafts or scenarios on various assumptions. The judge considered that it was not a requirement to produce a witness involved in the creation or keeping of the document to authenticate a document; other means of authentication may suffice. No evidence was tendered regarding the software used, the computer system in which it was stored, or the integrity of that system. Although the judge made the correct conclusion regarding the means of authentication, without evidence as to the integrity of the record keeping system, it could not be said that the court was satisfied with the means of authentication used.

1 (2005) 216 ALR 320, [2005] NSWSC 417.

7.44 Secondly, the defendants emphasized the importance of the documents to ASIC's case, and the fact that this was a civil penalty proceeding in which allegations were being made of serious misconduct.[1] However, the judge did not consider that

this submission affected the question as to whether documents had been adequately authenticated. Rather, documents can be authenticated by such evidence about their nature and provenance as will give rise to the inference that they were what ASIC claimed they were. Once they are adduced in evidence, it is open to the defendants to show that they have no probative value, for example by establishing that they are drafts not acted upon, or that they are based on assumptions or scenarios not widely held within the company. Austin J tempered this by saying that the law does not overload the authenticity requirement by including within it an obligation for the tendering party to rebut all such possibilities, and issues going to the ultimate probative value of the documents could not be assessed at that stage, because they did not bear on authentication. Austin J held that it would be setting the standard of authentication at too high a level to require ASIC to show, in the case of each document, that it is unique and not simply one of several versions. The point is that the issues regarding the authentication of electronic evidence *are* about the system in which the evidence, be they drafts or final versions, is what the court needs to consider.

1 Giving rise to the considerations enunciated by the High Court in *Briginshaw v Briginshaw* (1938) 60 CLR 336.

7.45 Austin J considered that it would be absurd, according to Bryson J in *NAB v Rusu*, for the law to dispense on a general basis with the need to prove the authenticity of a document:

> … for that would 'put the court entirely in the hands of whatever a document which a party chose to tender purported to be, subject to whatever opportunity another party had of overcoming its apparent effect'. On the other hand, it is important not to set the bar too high for the authentication of documents, because if too much is demanded, the authentication requirement will fight against the policy underlying the business records provisions which, as Hope JA remarked in *Albrighton* (at 548), is 'of great importance in the search for truth'. That policy recognises that any significant organisation depends for its efficiency upon the keeping of proper records, to be used and relied upon in the everyday carrying on of the activities of the business and therefore likely to be accurate, and 'likely to be a far more reliable source of truth than memory': *Albrighton*, at 548–9 per Hope JA; see also Australian Law Reform Commission, *Interim Report on Evidence*, Report No 26, vol 1, at [709]. It is reflected in the terms of s 69, which makes hearsay representations in business records admissible without requiring evidence from their authors.[1]

1 (2005) 216 ALR 320, [116].

7.46 Austin J concluded that in *NAB v Rusu*, Bryson J did not deny that inferences may be drawn from the document itself, relevant to the question of authenticity.[1] Austin J noted that apart from s 58(1), there is express statutory authority to do so in s 183, when a question arises in regards to the applicability of a provision of the Evidence Act. However, Austin J considered that *NAB v Rusu* insists on the need for authenticity to be established, and asserted that authentication cannot be achieved solely by drawing inferences from the face of the document where there is no other evidence to indicate provenance. In his opinion, the other cases do not deny these propositions.

1 (2005) 216 ALR 320, [117].

7.47 Austin J concluded that authentication is about showing that the document is what it is claimed to be, not about assessing the document. At the point of adducing the evidence, authentication is based on whether the document proves what the tendering party claims it proves. This means that a tendering party must show something more than the mere tender of the document itself when the tender is contested. If the tendering party adduces provenance evidence, then the court can conclude on the balance of probabilities that the document has been adequately authenticated. However, such evidence does not show who created the document, how the document was used within the organization, or even whether it was the only version, or whether it might have been a draft.

7.48 In *ASIC v Rich*, Austin J held that the question whether a particular document is one of several versions should be addressed in light of all the evidence, including such evidence as the defendants may choose to adduce. In that case, the documents in question were trial balances, and were all headed 'trial balances', and the end-of-month dates were specified. They were located in the finance directory of the I:/Drive, and their file paths indicate that they were trial balance or monthly balance sheet documents. The judge considered that the fact that no trial balances were tendered for July, August and October 2000 did not bear on the authenticity of the documents. Austin J stated that the fact that a trial balance, which was really for the month of April 2000, was incorrectly labelled 31 March 2000, did not prevent ASIC from authenticating the document. This was a matter to be decided once all the evidence had been adduced. The defendants set out a table of the trial balances, comparing asserted dates with dates 'modified' for documents where the document properties were available. The 'modified' dates were later than the asserted dates, and the defendants submitted that the court could not confidently draw an inference that the document in its tendered form was available within One.Tel at any particular time. In some cases the 'modified' date was well before the appointment of voluntary administrators and in other cases the modified date was at a crucial time, but the judge said these were matters going to probative value rather authentication.

7.49 There were also management accounts, all headed as such for the specified months. These documents, on their face, purported to be either profit and loss statements or statements of operating expenses. Some had footers indicating their character as management accounts and were located in the finance directory of the I:/Drive, the file paths of which also indicated their character as management accounts. The judge noted that the 'modified' dates for these documents were later than the asserted dates. In the absence of any appropriately qualified technical witness to give evidence on this point, the judge said they were anomalies which, if they were not explained by other evidence, would affect and possibly destroy the probative value of the documents in question. However, these facts did not go to the authenticity of the documents. It was the metadata within documents that could point to the integrity of documents produced. Without reverting back to the original software that generated the reports in question, such reports could be authenticated and relied upon. Most financial systems are contained within specially designed financial management software and reports, usually as spreadsheet programs, and are exported from the financial management software. There did not appear to have been any evidence about the financial management software, how the reports were generated, or evidence that the information that had been printed off was accurate.[1] While the judge

correctly stated that accuracy goes to probative value, and that it is for the other party to challenge the evidence, there did not appear to have been a challenge about the systems used to record, store and calculate financial reports of the company with the consequence that an old rule is being applied to new types of evidence.

1 For the importance of spreadsheet programs in the financial sector, see Grenville J Croll, 'The Importance and Criticality of Spreadsheets in the City of London', available at <www.eusprig.org/conference-abstracts.htm>. See also the discussion of spreadsheet programs below.

7.50 *ASIC v Rich* was applied in *Australian Competition and Consumer Commission v Allphones Retail Ply Ltd (No 4)*.[1] These cases apply an old rule of authentication to new types of evidence, that is, rules developed for paper are being applied to digital evidence. In *Australian Competition and Consumer Commission v Air New Zealand Limited (No 1)*,[2] the Federal Court of Australia stated that if there is an issue regarding the authenticity of a document, it may still be admissible if it is relevant or arguably so. This is provided there is material from which its authenticity may reasonably be inferred. That material will include what may reasonably be inferred from the document itself. Evidence about the system, its integrity, and how the reports were generated should have been tendered and competent witnesses should have given evidence as to the systems operation and reasonable level of security. An argument about provenance cannot be correctly posed and answered without such evidence.[3]

1 (2011) 280 ALR 97.
2 (2012) 301 ALR 326.
3 Steven W Teppler refers to 'testable reliability' in 'Testable Reliability' 255.

Other methods of authentication

7.51 It is not always necessary to obtain intricate details of a computer or its operating system before electronic evidence may be accepted into evidence, and the means by which a document is authenticated may not necessarily require the evidence of a suitably qualified expert, as in the case of *DPP v Brian Meehan*.[1] In this case, the members of the Republic of Ireland Court of Appeal were satisfied that evidence in digital form in relation to the records of telephone calls made and received was authenticated by appropriate witnesses. Kearns J said:

> When the telephone numbers on the computer printout were checked against the names who had registered each of the telephone numbers the identity of the users of each of the mobile phones was established clearly from the direct evidence which had been given by the various witnesses identified by the court.

1 [2006] IECCA 104, [2006] 3 IR 468.

7.52 Another example is the Canadian case of *Animal Welfare International Inc v W3 International Media Ltd*.[1] Ross J said that assessing the admissibility of electronic evidence comprises the following:

> The question at the admissibility stage does not involve an assessment of the likelihood that the evidence is accurate or true. The Supreme Court of Canada has emphasized that the reliability component of the principled approach involves a much lower standard of 'threshold' reliability, as distinct from the 'ultimate' reliability, the latter of which relates to the amount of weight to be accorded the evidence once admitted.[2]

1 2013 BCSC 2193.
2 2013 BCSC 2193, [64].

7.53 There is a great deal of misunderstanding, for instance, over the admissibility of email.[1] It is often asserted that because emails can be easily forged,[2] it is important to prove an email has not been forged before it is admitted into evidence. This proposition is not correct. Documents typed on paper may be forged[3] and altered, as in the case of *Scholastic, Inc. v Stouffer*[4] and letters may also be forged, as in the case of *Arrow Nominees, Inc v Blackledge*.[5] The forgery of evidence is nothing new, and just because it is possible to forge an email, it does not mean that email correspondence is required to undergo an extensive forensic analysis to prove it is not a forgery for it to be admitted into evidence.[6] The authenticity of a document in digital form can be tested in other ways that are equally as effective. For instance, in *R v Boulkhrif*,[7] the defence objected to the reliability and accuracy of bank transfers recorded on what appeared to be computer print-outs. The documents included the initials of a bank clerk. The members of the Court of Appeal indicated that the initials provided evidence that transfers were authorized, but the presence of the initials did not prove the authenticity of the document. In this case, this particular ground of appeal was made out, because if the document tendered was a computer print-out, there was insufficient evidence of authentication, although in his commentary, Professor Smith pointed out that the purpose of the initials was not clear. This example illustrates the difference between the accuracy of the content recorded on the print-out and the authenticity of the print-out. At times, the term 'accuracy' might be understood in the sense of 'authenticity', but it is always possible to dispute the facts recorded in the authenticated print-out.[8]

1 Early employment law cases in England in which emails featured (some where email was the main evidence) prominently include: *Pennington & Beverly v Holset Engineering Limited* (30 August and 7 November 2000, unreported) (Case Nos 1802184/00 and 1802185/00), Leeds Employment Tribunal (dismissed for circulating offensive emails); *Bower v Schroder Securities Limited* (Hearings throughout 2000, 2001 and 2002, unreported) (Case Nos 3203104/99 and 3203104/99/S), London Central Employment Tribunal (unfairly dismissed, unlawful sex discrimination and equal pay); *Royal Bank of Scotland v Goudie* (Appeal No. UKEAT/0693/03/TM) (unfairly dismissed for sending emails containing pornographic content);*Villalba v Merrill Lynch & Co Inc, Merrill Lynch Europe Limited and Merrill Lynch International Bank Limited* (2003, unreported) (Case Nos 2302467/2003 and 2305203/2003) (UKEAT/0461/04/TM, UKEAT/0223/05/LA) (unlawful victimization and unfair dismissal); *Crook v Manpower plc* (30 May 2001, unreported) (Case No 1501774/2000), Bury St Edmunds Employment Tribunal (fair dismissal because of content of an email about a female member of staff); *Jayyosi v Daimler Chrysler Limited* (Hearings in February and March 2003, unreported) (Case No 1201592/02), Bedford Employment Tribunal (unlawful acts of racial discrimination).

2 An email was forged and presented to the court as legitimate evidence in the case of *Munshani v Signal Lake Venture Fund II, LP*, 805 N.E.2d 998 (Mass.App.Ct. 2004); 2001 WL 1526954 (Mass.Super.), which led to Munshani being indicted on counts of criminal attempt and obstruction of justice; in *Greene v Associated Newspapers Limited* [2004] EWCA Civ 1462, the technical evidence that emails were forgeries was not clear.

3 Winsor C Moore, 'The questioned typewritten document' (1959) 43 Minn L Rev (1959) 727.

4 221 F.Supp.2d 425 (S.D.N.Y. 2002). See *Breezevale Limited v Dickinson*, 879 A.2d 957 (D.C. 2005) for submission of a range of forged evidence, including computer documents produced before Breezevale even had computers; see also *Masood v Zahoor* [2008] EWHC 1034 (Ch); *Zahoor v Masood* [2009] EWCA Civ 650 where the trial judge reached the conclusion that both parties committed forgery and perjury.

5 [2000] All ER (D) 854; [2000] 2 BCLC 167; [2001] BCC 591 reversing [1999] All ER (D) 1200; [2000] 1 BCLC 709.

6 Emails can, of course, be forged: *R v Debnath* [2005] EWCA Crim 3472. See also report of a prisoner escaping Wandsworth prison by forging an email saying he had been granted bail: 'Wandsworth Prison escapee Neil Moore faked bail email' (*BBC News*, 6 August 2016) <www.bbc.com/news/uk-england-london-32095189>.
7 [1999] Crim LR 73.
8 Seng, 'Computer output as evidence', 164.

7.54 In comparison, in *R v Mawji (Rizwan)*,[1] the appellant was convicted of making a threat to kill, and part of the evidence included an email sent to the victim dated 31 July 2002, which read:

> Hi Bitch,
>
> Don't think you're safe in the UK. I'm going to kill you.
>
> I will make sure I get my hands on you … waiting for you.
>
> Your loving husband.
>
> Riz.

1 [2003] EWCA Crim 3067, [2003] All ER (D) 285 (Oct).

7.55 A witness for the defence gave evidence to demonstrate how relatively easy it was to produce a document that claimed to be an email, but which had nothing to do with the email account from which it purported to come. It was suggested that somebody else was responsible for sending the email in question. One of the grounds of appeal was that the email was a copy and was secondary evidence (which is correct) if adduced in the form of a print-out, and it was necessary to provide evidence of the audit trail or some other similar evidence to show the authenticity of the document. The members of the Court of Appeal rightly rejected this submission. The analysis offered by Kay LJ centred upon the evidence produced by the victim when she saw the email on the screen and then printed it. However, the email did not have to be authenticated in the way suggested by the appellant because of the circumstances surrounding the events and the other evidence in the case. The content of the email was similar to other evidence produced at trial, which went to show that the email was written and sent by the appellant, and the members of the jury had to consider whether, in all the circumstances, it was possible that somebody else might have produced the email. Hence the content of the email demonstrated its authenticity on the face of the totality of the evidence. If the email was fabricated, it had to be questioned as to why somebody would go to the length of forging the content of an email that was so obviously linked to the other evidence, which was indisputably written by the appellant, produced at trial.[1]

1 In *Clifford v The Chief Constable of the Hertfordshire Constabulary* [2008] EWHC 3154 (QB), Mr Justice Cranston observed, at [27] that 'Where images are of a type generated from a site which the user has previously purchased materials of a similar nature, that may give rise to an inference that the user knows they are stored in the computer.' For a similar analysis in the context of Canada, see David M Paciocco, 'Proof and Progress: Coping with the Law of Evidence in a Technological Age' (2013) 11 Canadian Journal of Law and Technology 181.

7.56 The use of metadata in an email header can prove that an email was sent and received, and therefore go to show that the email is not a forgery. In *Greene v Associated Newspapers*,[1] an email was analysed to establish its authenticity. Emails were alleged to have been exchanged between Peter Foster and Martha Greene, a close friend of Cherie Blair. Ms Greene denied sending the emails to Peter Foster and claimed the emails

were forgeries. A digital evidence professional who examined Ms Greene's computers could find no trace of the emails. Another digital evidence professional inspected three emails on a laptop owned by Mr Foster at his home in Australia. This specialist was able to complete a 'trace route' on the IP address headers. This evidence was sufficient to indicate that the emails were sent from a server in the Greater London area.[2] The mail servers that were reported in the email header were actual servers, and the times recorded by the email header indicated that the times received were accurate. The email address header from the sender could not be changed, although the sender of the email could have been another person who had access to the owner's computer. Upon inspection of the email header, it showed that from the point of departure to the addressee's inbox, the emails had not been interfered with. The defendant's digital evidence professional stated that although the text of an email could be altered upon forwarding or sending the email to oneself or to a third party, the original header would reflect this change, and there was no such indication in the header information in the emails in question. The Court of Appeal agreed with the trial judge that there was no clear 'knock-out' evidence to show the email was a forgery.[3] The same arguments were used in the US case of *People of the State of Illinois v Downin*,[4] where a digital evidence professional testified that the only way of authenticating the origin of one of the emails in question was by investigating the IP address,[5] which was not included on the exhibit. As in the case of *R v Mawji (Rizwan)*, the relevant email contained admissions of guilt. However, it was not necessary to authenticate the email by providing evidence of the IP address and then linking the continuity of evidence to the sender of the email, as pointed out by O'Brien J:

> A finding of authentication is merely a finding that there is sufficient evidence to justify presentation of the offered evidence to the trier of fact and does not preclude the opponent from contesting the genuineness of the writing after the basic authentication requirements are satisfied. ... The prosecution need only prove a rational basis upon which the fact finder may conclude that the exhibit did in fact belong to the defendant.[6]

1 *Greene v Associated Newspapers* [2004] EWCA Civ 1462.
2 [2004] EWCA Civ 1462, [37].
3 [2004] EWCA Civ 1462, [21].
4 357 Ill.App.3d 193, 828 N.E.2D 341 (Ill.App. 3 Dist. 2005).
5 Birss QC J pointed out, at [28] in *Media CAT Limited v Adams* [2011] EWPCC 6, that 'All the IP address identifies is an internet connection, which is likely today to be a wireless home broadband router. All Media CAT's monitoring can identify is the person who has the contract with their ISP to have internet access'.
6 828 N.E.2d 341 (Ill.App. 3 Dist. 2005), 350.

7.57 A digital document may be authenticated by direct or circumstantial evidence, and circumstantial evidence includes a range of factors, including, but not limited to, appearance and the contents of the document, the subject matter, witness testimony, and any distinctive features that indicate a nexus, as demonstrated in the case of *United States of America v Simpson*,[1] where the United States Court of Appeals for the Tenth Circuit concluded that the circumstantial evidence produced by the government demonstrated that the appellant was in possession of abusive images of children. Where the content demonstrates knowledge of the circumstances of the facts such that only very few people in the world will be aware of them, the inference as to authenticity must be overwhelming in most cases.[2] Proving of the authenticity of instant messages

is also perfectly possible through the use of compelling circumstantial evidence. In the case of *In the interest of F.P., a minor*,[3] Ford Elliott J offered some robust and realistic comments on this topic that bear repeating:

> Essentially, appellant would have us create a whole new body of law just to deal with e-mails or instant messages. The argument is that e-mails or text messages are inherently unreliable because of their relative anonymity and the fact that while an electronic message can be traced to a particular computer, it can rarely be connected to a specific author with any certainty. Unless the purported author is actually witnessed sending the e-mail, there is always the possibility it is not from whom it claims. As appellant correctly points out, anybody with the right password can gain access to another's e-mail account and send a message ostensibly from that person. However, the same uncertainties exist with traditional written documents. A signature can be forged; a letter can be typed on another's typewriter; distinct letterhead stationery can be copied or stolen. We believe that e-mail messages and similar forms of electronic communication can be properly authenticated within the existing framework. … We see no justification for constructing unique rules for admissibility of electronic communications such as instant messages; they are to be evaluated on a case-by-case basis as any other document to determine whether or not there has been an adequate foundation showing of their relevance and authenticity.[4]

1 152 F.3d 1241, 1249 (10th Cir. 1998).

2 As in *Dickens v State of Maryland*, 175 Md.App. 231, 927 A.2d 32 regarding text messages sent and received over mobile telephones; see *Commonwealth of Pennsylvania v Koch*, 2011 WL 4336634 (Pa. Super.), 2011 P.A.Super 201 where the evidence was not sufficient to authenticate text messages.

3 *In the interest of F.P., a minor*, 878 A.2d 91 (Pa.Super. 2005), 2005 PA Super 220.

4 878 A.2d 91 (Pa.Super. 2005), 95; this case was cited favourably in *State of North Dakota v Thompson*, 777 N.W.2d 617 (N.D. 2010), 2010 ND 10 ; this case was cited favourably in *State of North Dakota v Thompson*, 777 N.W.2d 617 (N.D. 2010), 2010 ND 10.

7.58 The judge listed a number of cases involving the authentication of email communications,[1] web pages[2] and chat room exchanges,[3] all of which included circumstantial evidence that provided compelling evidence to prove the authenticity of the documents in question.[4]

1 *Massimo v The State of Texas*, 144 S.W.3d 210 (Tex.App.-Fort Worth 2004); *Kearley v State of Mississippi*, 843 So.2d 66 (Miss.App. 2002), certiorari denied, 842 So.2d 579 (Miss. 2003); *United States v Siddiqui*, 235 F.3d 1318 (11th Cir. Ala. 2000), certiorari denied 533 U.S. 940, 150 L.Ed.2d 737, 121 S.Ct. 2573 (2001).

2 *Perfect 10, Inc. v Cybernet Ventures, Inc*., 213 F.Supp.2d 1146 (C.D.Cal. 2002); see also *Zhu v Merrill Lynch HSBC* 2002 BCPC 0535 where the print-outs of screen shots differed as between the parties, *Wady v Provident Life and Accident Insurance Company of America*, 216 F.Supp.2d 1060 (C.D.Cal. 2002) where documents downloaded from a website were not authenticated satisfactorily; and *Hutchens v Hutchens-Collins*, 2006 WL 3490999 (D.Or.), where documents found on a website connected to one of the parties and not protected by a password were sufficiently authenticated.

3 *United States of America v Tank*, 200 F.3d 627 (9th Cir. 2000).

4 Articles on this topic, with extensive references to relevant case law in the US include: Catherine Guthrei and Brittan Mitchell, 'The Swinton Six: The impact of State v. Swinton on the authentication of digital images' (2007) 36 Stetson L Rev 661; Cooper Offenbecher, 'Admitting computer record evidence after In Re Vinhnee: A stricter standard for the future?' (2007) 4 Shidler J L Com & Tech 6; Steven Goode, 'The admissibility of electronic evidence' (2009) 29 Rev Litig 1; Jonathan D Frieden and Leigh M Murray, 'The admissibility of electronic evidence under the Federal Rules of Evidence' (2011) 17 Rich J L & Tech 5; Kenneth N Rashbaum, Matthew F Knouff and Dominique Murray, 'Admissibility of non-U.S. electronic evidence' (2012) 18 Rich J L & Tech 9.

7.59 In circumstances where there is no evidence from a digital evidence professional, the adjudicator has no other option other than to assess disputed evidence of data in digital form by evaluating the evidence of witnesses,[1] establishing what their respective technical skills with computers might be as revealed at trial, and the probability of whether emails, text messages and such like were sent as alleged. Mr Justice Dingemans had to assess electronic evidence in the case of *GH Cornish LLP v Smith*[2] without the aid of evidence from a digital evidence professional. Neither party adduced any expert forensic evidence, although much of the evidence comprised emails, text messages and data from websites. GH Cornish did call a digital evidence professional, but the witness did not give evidence as an expert, but as a witness of fact. The judge commented that:

> Given the apparent absence of a good reason for calling Miss Collie as an expert witness, and the absence of an explanation for the failure to pursue the claim for trespass to the computers by installing 'malware', I have formed the distinct impression, and find, that I have not been told all that I could be about the forensic investigations which were carried out. I accept it is for a party to adduce the evidence that it chooses to prove its case, but in circumstances where there is no direct evidence, gaps in the evidence make the drawing of reliable inferences more difficult.[3]

1 As in the Canadian case of *R v Galuce Nde Soh* 2014 NBQB 020, where LaVigne J held a trial within a trial to determine the status of digital data comprising the exchange of messages from a social networking site.
2 [2013] EWHC 3563 (QB).
3 [2013] EWHC 3563 (QB), [20].

7.60 In reaching his conclusions, Mr Justice Dingemans noted the following when assessing the evidence:

> (i) There was no forensic evidence to link Mr Smith with the disputed materials. The absence of forensic evidence was not fatal to the claim, which can be proved in a number of different ways. However, the absence of such evidence was relevant in circumstances where Mr Smith was, as the judge found, naïve about what could be located about his actions on his own work computer and on the Internet.[1]
>
> (ii) The judge accepted that there was a coincidence of timing and interest in the subject matter of many of the disputed materials, which supported the claimants' belief that Mr Smith was the author of the disputed materials. However, the evidence also showed that another person would have known about the matters, and in such circumstances it is difficult to see how such a person could be excluded from responsibility for the materials.[2]
>
> (iii) The judge had not been told all that he could have been told about forensic investigations, which meant he was careful about drawing inferences because of the absence of evidence. He assessed the evidence which was before him as it was.[3]

1 [2013] EWHC 3563 (QB), [141].
2 [2013] EWHC 3563 (QB), [142].
3 [2013] EWHC 3563 (QB), [145] and [146].

7.61 The authenticity of pages from the Internet can be of some concern, however, because they can alter frequently. For instance, in *R v Skinner (Philip)*,[1] there was no reliable foundation evidence for the introduction of screen images, as in the Canadian case of *Jalil v Canada (Minister of Citizenship and Immigration)*[2] where an immigration

officer relied on two documents posted on websites, the contents of which were successfully challenged. Mosley J commented that: 'The integrity of the process of determining whether there are reasonable grounds to believe that an individual is a member of an organization that has engaged in terrorist activities deserves greater diligence than was displayed in this instance'.[3] The time and date that a person adjusts the information contained on a website might also act to exonerate a person accused of a crime. Rodney Bradford was arrested and held in custody for 12 days in connection with an armed robbery of two people in the Brooklyn housing project where he lived. He claimed that he was physically in Manhattan at the time the crime took place, and his alibi was by way of an update he made to his page on a social networking website from a computer located in his father's home in Manhattan. The office of the District Attorney acknowledged that this evidence helped to dismiss the charges against him.[4]

1 [2005] EWCA Crim 1439, [2005] All ER (D) 324 (May), [2006] Crim LR 56.
2 2006 FC 246.
3 2006 FC 246, [40].
4 Edith Honan, 'Facebook provides alibi for robbery suspect' (Reuters, 12 November 2009) <http://uk.reuters.com/article/2009/11/12/us-facebook-alibi-idUSTRE5AB5JO20091112>.

7.62 In comparison, websites of government departments are considered to be self-authenticating in the US, as in *Williams v Long*,[1] where print-outs from the Maryland Judiciary Case Search website and Employment Standards website were held to be self-authenticating because the websites were publications of a public authority. Similarly, in *Paralyzed Veterans of America v McPherson*,[2] the plaintiffs' relied on a number of materials that appeared on the Secretary of State's website, the first of which comprised two documents: a letter dated 6 December 2007 approving the use of the AutoMARK model 200 for the February 2008 election, and the second, a set of conditions on the use of the AutoMARK for all California counties using this voting machine, issued on the same day. The court was requested to take judicial notice of these documents, pursuant to Federal Rules of Evidence 201(b)(2). The relevant rule reads as follows:

> A judicially noticed fact must be one not subject to reasonable dispute in that it is either (1) generally known within the territorial jurisdiction of the trial court or (2) capable of accurate and ready determination by resort to sources whose accuracy cannot reasonably be questioned.

1 585 F.Supp.2d 679 (D. Md. 2008).
2 2008 WL 4183981 (N.D. Cal. Sept. 9, 2008).

7.63 Armstrong DJ cited a number of cases regarding this topic, and granted the request, observing that:

> The documents being requested for judicial notice are not disputed and their accuracy is not reasonably questioned. Moreover, there is no objection by any of the defendants to the plaintiffs' first request for judicial notice. Accordingly, the Court will take judicial notice of the letter and the Secretary of State's expressed conditions on the use of the AutoMark.[1]

1 2008 WL 4183981 (N.D. Cal. Sept. 9, 2008), 5 and 6.

7.64 Regarding a second set of materials from the website, one objection was raised that some of the materials were not authenticated and there was no appropriate

foundation laid. This argument was rejected on the basis that the print-outs from the Secretary of State's website pages were official records for purposes of Rule 902(5), which reads:

> Rule 902. Evidence That Is Self-Authenticating
>
> The following items of evidence are self-authenticating; they require no extrinsic evidence of authenticity in order to be admitted:
>
> …
>
> (5) *Official Publications*. A book, pamphlet, or other publication purporting to be issued by a public authority.

7.65 This meant that the documents were therefore self-authenticating. The judge also cited a number of cases regarding the provisions of Rule 902(5) and how it applied to documents from government websites.[1]

1 2008 WL 4183981 (N.D. Cal. Sept. 9, 2008), 7.

The threshold for authentication

7.66 The threshold for authentication must be the same as for any other form of evidence: there must be prima facie evidence to support the claim.[1] In the Intellectual Property Office case of *HSBC France*,[2] HSBC France appealed against the decision of the patent examiner, who refused to grant a patent for authenticating users of a website. The application was rejected because of a previous patent filed by Fujitsu Services (published on 12 March 2003), and because of an article by Laurika Bretherton 'Banking on Trust' that appeared in *Computing* on 20 February 2004. The article included a case study from Lloyds TSB, which briefly described the bank's keystroke logging software that records what a user is typing on the keyboard, thus avoiding use of the keyboard altogether. If the date of the article was correct, it appeared five months before the priority date. HSBC appealed the ruling.

1 *R v Robson (Bernard Jack), R v Harris (Gordon Federick)* [1972] 2 All ER 699, [1972] 1 WLR 651, CCC per Shaw J at 654 E-F.
2 BLO/180/09 (29 June 2009), Hearing Officer B Micklewright.

7.67 HSBC disputed the reliability of the date of the article. The examiner was not able to locate a paper copy of the relevant issue, and the only evidence was the date specified on the website. The page had not been archived. The Hearing Officer indicated that there were two issues: first, whether the article was made available to the public on 20 February 2004, and second, whether it had been altered since its initial publication. HSBC argued that the Hearing Officer should follow the reasoning adopted in the European Patent Office Technical Board of Appeal decision in *Konami Limited* T 1134/06 in relation to the reliability of dates of web pages. In this case, the Board decided that 'the criteria to be applied for establishing a disclosure made available to the public through the internet as in the present case should be the same as those introduced by the jurisprudence of the Boards of Appeal for establishing a prior use or a prior oral disclosure … These questions are to be decided using the same strict standard of proof in respect of a prior use or a prior oral disclosure.'[1] The standard of proof was to be beyond reasonable proof,[2] although the Board accepted that there might be occasions when other evidence might be helpful in assessing the authenticity of the document:

> 4.2 In certain cases, where a website belonging to a reputable or trusted publisher publishes online electronic versions of paper publications, content and date can be taken at face value, and the need for supporting evidence can be dispensed with. It can also be envisaged that if a web site operates under recognized regulations and standards, which allow date and content of information retrieved therefrom to be established with a high degree of certainty, further evidence may also not or no longer be required. Of course, it should be clear for both the examiner and the public whether an internet source is considered as 'reputable' or 'regulated'. This again calls for clearly defined guidelines.
>
> Where a disclosure has been retrieved from a resource such as the Internet Archive, further evidence concerning the history of the disclosure, whether and how it has been modified since the date it originally appeared on a web site will be necessary. This could be in the form of an authoritative statement from the archivist. Alternatively, an appropriate statement as to the content, either from the owner or author of the archived web site which included the disclosure may suffice.

1 At 2.4.
2 At 4.1.

7.68 The Hearing Officer pointed out that the law in England and Wales took a somewhat different approach to the burden of proof required in establishing publication dates for documents and dates for cases of alleged prior use:

> This is summarised very clearly in paragraph 2.29.1 of the Manual of Patent Practice, which states:
>
> '2.29.1 In cases of alleged prior use, the required standard of proof is the balance of probabilities. Within this standard, the Patents County Court in *Kavanagh Balloons Pty Ltd v Cameron Balloons Ltd* [2004] RPC 5 held that a flexible degree of probability should be applied to evidence relating to prior use. The cogency of the evidence had to match the occasion and be proportionate to the subject matter. Because of the nature of the monopoly itself and question of public interest, no stricter standard should be applied. It was held that it was not necessary for an opponent to prove his case "up to the hilt" as had been required by the EPO Technical Board of Appeal in *Sekisui/shrinkable sheet* [1998] OJEPO 161 (T 472/92). The hearing officer in *Colley's Application* [1999] RPC 97 also distinguished from Sekisui by not requiring proof "up to the hilt", but followed this decision and *Demmeler Maschinenbau GmbH & Co KG* (T 908/95) in holding that mere assertion of prior use was insufficient: place, time and detail were essential.'[1]

1 BLO/180/09 (29 June 2009), 17.

7.69 The practice in the UK differed from the European Patent Office, and the Hearing Officer did not feel bound to follow the decision in *Konami Limited* T 1134/06. He determined that the article would be assessed on the balance of probabilities. In reaching his decision, the Hearing Officer also referred to guidance provided by the Manual of Patent Practice:

> Other relevant guidance on this matter is found in paragraphs 18.09.2 and 18.09.3 of the Manual of Patent Practice which state:
>
> > '18.09.2 As Mann J indicated in *Macrossan's Patent Application* [2006] EWHC 705 (Ch), any doubt should be resolved in the applicant's favour only if the doubt is substantial. This could arise if the examiner's assertions as to the common general knowledge have been challenged

> and expert evidence would be needed to establish the position, or if the date of a prior disclosure has been challenged and the examiner does not have access to material that would confirm the date. Certainly the examiner is not required to meet the criminal burden-of-proof standard in raising and pursuing an objection.
>
> 18.09.3 When assessing the relevance of an internet disclosure at the substantive examination stage, a document should be cited unless the examiner is certain that it falls outside the state of the art. If the applicant contests the publication date the examiner should decide the matter on the balance of the evidence available. Evidence from sources such as archive.org, while not conclusive, may provide justification for an examiner's view that there is little doubt as to the date of disclosure.'
>
> These paragraphs support my view that I should decide the matter of a contested date of an internet disclosure on the balance of the evidence available. Only if there is substantial doubt should I give the applicant any benefit of the doubt.[1]

1 BLO/180/09 (29 June 2009), 19.

7.70 The Hearing Officer assessed the evidence of the article and concluded that the date of the publication was demonstrated. He considered that Incisive Media Ltd, which also published magazines, operated the website, together with a second website. The publisher provided information to businesses in relation to IT matters, and was well known for so doing. Each article that it published included a date. Although the examiner was not able to obtain a paper copy of the magazine in which the article appeared to be printed, nevertheless the Hearing Officer considered it highly likely that the article did appear in a paper version of *Computing* magazine on or around 20 February 2004. Even if this was not the case, the Hearing Officer noted that the websites were highly reputable and had a long history of publishing for both IT consumers and the IT industry. Taking this into account, he considered that the dates given to articles were reliable, and that the content of the articles were unlikely to be altered once they were published. The Hearing Officer concluded that on the balance of probabilities the Internet disclosure was made available to the public on 20 February 2004. He then went on to state:

> In fact, even if I was to follow the reasoning in T 1134/06, I would have concluded that the date of the internet article had been proved 'up to the hilt' and could be taken at face value without the need for supporting evidence given the trusted and reliable nature of the websites in question and their links with paper publications.[1]

1 BLO/180/09 (29 June 2009), 21.

Proof of authentication as a matter of law

7.71 There will be occasions when it will be necessary to provide evidence that a digital document is authentic as a matter of law, and circumstantial evidence will not be sufficient, as in the Australian case of *Roads and Traffic Authority of New South Wales v Timothy Adam Michell*.[1] In this case, a speed measuring device took a photograph of a vehicle being driven by the defendant, which demonstrated he was driving at a speed greater than the speed limit in force. At the trial, the defence objected to the photograph being submitted in evidence because the security indicator, which every such photograph had to bear as prescribed by regulations, had been struck through.

The prosecution did not rely upon the security number that was struck through, and the magistrate ruled that the photograph was admissible in evidence. At the end of the prosecution's case, the magistrate ruled there was a case to answer. Mr Michelle neither gave nor called any evidence, and the magistrate then had to decide whether he was satisfied beyond reasonable doubt that the offence had been proved. He concluded that he could not be satisfied that the photograph had not been altered since it was taken, and he therefore acquitted Mr Michell. On appeal, Adams J agreed that the magistrate was correct in admitting the photograph in evidence. In dealing with the security number, the judge was required to interpret the provisions of s47 of the Road Transport (Safety and Traffic Management) Act 1999 (NSW). The relevant provisions are as follows:

> 47 Photographic evidence of speeding offences
>
> (1) In proceedings for an offence of driving at a speed in excess of a speed limit imposed by or under this Act or the regulations, evidence may be given of a measurement of speed obtained by the use of an approved speed measuring device and recorded by an approved camera recording device.
>
> (2) In proceedings in which such evidence is given:
>
> (a) the provisions of section 46 relating to the accuracy or reliability of the approved speed measuring device apply, and
>
> (b) subsections (3)–(5) apply in relation to the approved camera recording device, and
>
> (c) evidence that a photograph taken by an approved digital camera recording device bears a security indicator of a kind prescribed by the regulations is evidence (unless evidence to the contrary is adduced) that the photograph has not been altered since it was taken.
>
> (3) A photograph tendered in evidence as a photograph taken by an approved camera recording device on a specified day at a specified location:
>
> (a) is to be accepted as having been so taken (unless evidence to the contrary is adduced), and
>
> (b) is evidence (unless evidence to the contrary is adduced) of the matters shown or recorded on the photograph.

1 [2006] NSWSC 194.

7.72 It was held that the true construction of s 47(2)(c) meant that the legislature deemed it necessary to provide for the authenticity of the photograph. Otherwise the provisions of the sub-section would be redundant. This therefore required the prosecution to adduce evidence of the authenticity of the photograph in order to demonstrate it had not been altered from the moment it was taken, to the moment in time it was used to demonstrate that an offence had taken place.

7.73 In the Australian case of *Alan Yazbek v Ghosn Yazbek*,[1] the court held that a will that had been created on the personal computer of the deceased was a will for the purposes of the Probate and Administration Act 1898 (NSW), and was admitted into probate. In that case, the court considered in some detail the expert evidence provided by a forensics specialist to show that the deceased had created his last will and testament on his computer. The judge considered the evidence of a computer forensic expert who demonstrated, by reviewing the metadata of the document named 'Will.doc' on the computer of the deceased, that the deceased had created the

document just prior to his death. Although there was a print-out of the document, the judge held that it was the electronic version of the document that was the deceased's last will and testament.

1 [2012] NSWSC 594 (1 June 2012).

7.74 Proving the case is paramount. The British Columbian case of *R v Nardi*[1] demonstrates that if evidence created by software programs is to be authenticated, sufficient evidence needs to be collected and provided to the court. In that case, the complainant had purchased a laptop computer and had downloaded 'undercover' security software, which purported to take screenshots and photographs of any user who operated the computer if it was stolen, identified the IP address being used, and provided a GPS coordinate of the location of the computer. It also allowed the computer to be disabled. The complainant's laptop was stolen and the Crown sought to tender the screenshots and photographs collected by the 'Undercover' software. The information had been collected before the complainant disabled the computer and printed the material for the police. As to whether the records were real evidence, the Crown failed to adduce any evidence in this regard.

1 2012 BCPC 0318.

7.75 The screenshots, photographs and location information that the complainant downloaded from his account with the provider did not contain any information as the computer being used. The Crown failed to file an affidavit from a representative of the provider showing that the software was functioning properly, or that the screenshots and photographs came from the complainant's computer. The judge rejected the submission that the documents were 'business records', because there was insufficient evidence to suggest that the records were kept in the usual and ordinary course of business or even if the provider's business was carried out in Canada. The Crown did not provide evidence on these points.

Considerations to be taken into account

7.76 The tests of authenticity for digital data (or perhaps digital objects or digital artefacts)[1] will vary, depending on the source and type of the data.[2] Lawyers must look to the digital forensic professionals for guidance. For instance, the print-out from a mainframe computer will demand a different approach in comparison to the data held on a personal computer; this in turn will be different if data is stored with a cloud service provider. The mainframe computer cannot be removed, so reliance must be placed on the print-outs and relevant expert evidence, which in turn raises the question of how is the reliability of the mainframe to be tested.

1 For an exploratory essay on 'data', see Lee A Bygrave, 'The meaning of "data" and similar concepts', in C Magnusson Sjöberg and P Wahlgren (eds), *Festskrift till Peter Seipel* (NorstedtsJuridik 2006).

2 For an interesting introduction, see Richard Boddington, Valerie Hobbs and Graham Mann, 'Validating digital evidence for legal argument', Proceedings of the 6th Australian Digital Forensics Conference, Edith Cowan University, Perth Western Australia, 3 December 2008, 41–58, available at <http://conferences.secau.org/proceedings/2008/forensics/ADFC%20Proceedings.pdf>.

7.77 In comparison, an image can be taken of the disk of a personal computer, and the data recorded on the copy can then be analysed without affecting the original disk. The imaging software will take a 'snapshot' of what the imaging software has found on the hard drive of the personal computer at that point in time, and digital evidence professional can testify as to the continuity of custody. However, whether files that were captured as part of the snapshot image were altered prior to the image being taken can be difficult to prove.[1]

1 The reader should be aware of the problems with anti-forensics and the limitations of imaging software, which is dealt with in more detail elsewhere in this text.

7.78 It is difficult to be clear as to what is meant by an 'authentic' digital object. If, for instance, a particular macro (for instance a macro that is used to automate frequently used movements of the mouse) is missing from a computer upon which a copy of the digital document rests, the question that must be raised is whether the lack of the macro in the computer in which the data now rest, renders the document something other than the genuine document. This blurs the distinction between what might be considered as the document and as the environment in which it exists.

7.79 To a certain extent, the technical focus of proving the authenticity of a digital object is to have checks and balances in place to demonstrate the history of how the data have been managed, which leads to the assertion that the data have not been modified, replaced or corrupted and must, therefore, be trustworthy. This proposition rests on two conditions: the data are subject to a continuity of custody; and the data have not been modified without authority between the time they were created or added to the depository, to the moment they were required. Jeff Rothenberg points to this problem:

> The first of these conditions is only a way of supplying indirect evidence for the second, which is the one that really matters. An unbroken chain of custodianship does not in itself prove that records have not been corrupted, whereas if we could prove that records had not been corrupted, there would be no logical need to establish that custodianship had been maintained. However, since it is difficult to obtain direct proof that records have not been corrupted, evidence of an unbroken chain of custodianship serves, at least for traditional records, as a surrogate for such proof.[1]

1 Jeff Rothenberg, 'Preserving authentic digital information', in *Authenticity in a Digital Environment* (Council on Library Information Resources 2000) 57.

7.80 This might be a discussion of the best evidence rule, but it is highly pertinent to the authentication of digital data.

Authentication and the best evidence rule

7.81 The unique nature of digital data means that although the data may be created in program memory, they might be saved on a number of different storage media formatted to different specifications. For instance, if a file is copied from a storage device formatted with Microsoft NTFS to one formatted with Solaris ZFS, the representation of the file on the media will be significantly different, in the encoding of both the data (especially if compression or encryption is enabled in the file system) and metadata. Further, each digital object may be replicated in a number of places, which means

that there is no single 'original', as it happens when disk arrays are aggregated into resilient file system volumes using a Redundant Array of Independent Discs (RAID) management. This has implications for understanding the nature of digital data. In essence, there is a need to accept that the concept of an 'original' and 'authentic' digital object is meaningless.[1] Therefore it is necessary to consider the meaning of 'authentic' in terms of a digital object in the relevant context. Rothenberg proposes the following:

> … we therefore define a digital original as any representation of a digital informational entity that has the maximum possible likelihood of retaining all meaningful and relevant aspects of the entity.
>
> This definition does not imply a single, unique *digital-original* for a given digital informational entity. All equivalent digital representations that share the defining property of having the maximum likelihood of retaining all meaningful and relevant aspects of the entity can equally be considered digital-originals of that entity. This lack of uniqueness implies that a digital-original of a given entity (not just a copy) may occur in multiple collections and contexts. This appears to be an inescapable aspect of digital informational entities and is analogous to the traditional case of a book that is an instance of a given edition: it is an original but not the original, since no single, unique original exists. (Italics in the original)[2]

1 George L Paul, *Foundations of Digital Evidence* (American Bar Association 2008), 48–9; Stephen Mason, 'Electronic evidence and the meaning of "original"' (2009) 79 Amicus Curiae 26, available at <http://sas-space.sas.ac.uk/2565/>.
2 Jeff Rothenberg, 'Preserving authentic digital information' 66.

Authenticating a digital object

7.82 As mentioned earlier, the concept of an 'original' document no longer exists with digital evidence.[1] Rather, a digital object is authenticated by verifying the claims that are associated with the object, such as:

> 1. The organizational criteria demonstrating the provenance of the digital object, including the documentation pertaining to the continuity of custody (and the extent to this documentation can be trusted), and the extent to which the custodians can be trusted.
>
> 2. The object can be examined forensically to establish whether its characteristics and content are consistent with the claims made about it and the record of its provenance, although the methods used may also be subject to challenge – for instance, how a mainframe is tested for reliability or consistency of output.
>
> 3. Whether the imaging techniques are challenged when relying on the evidence from a personal computer.
>
> 4. Any signatures, seals and time stamps[2] that may be attached to the object which can help test the claims to consistency and provenance.

1 See Chapter 3 for a more comprehensive discussion on this point.
2 Malcolm W Stevens, 'Unification of relative time frames for digital forensics' (2004) 1 Digital Investigation 225.

7.83 Luciana Duranti and Corinne Rogers frame the issues regarding the authentication of digital data as follows:[1]

> Can the data be trusted? Can the records from which the data are derived to be trusted? Are these records complete? Are they authentic? How were they generated, by whom and under what conditions? Is there sufficient contextual information to enable them to be understood?

1 Duranti and Rogers, 'Trust in digital records' 522, 523.

7.84 Proving the provenance of a digital object can be difficult, especially if files are copied and re-worked.[1] For instance, time stamps are used to indicate when a digital object was written, but the time stamp might not be accurate, or it might have been tampered with. Similarly, digital signatures are also used for the same purpose, but any confidence in the integrity (meaning that the data have not been corrupted) of the object can only be as good as the authenticity and integrity of the hash digest and the processes and procedures surrounding the digital signature and the equipment which generates and applies it.[2] It can be argued that the process of demonstrating the authenticity of a digital object is 'a process of examining and assigning confidence to a collection of claims.'[3] In essence, the ability to prove the authenticity of a digital object is not proving that an 'original' exists, especially when referring to something as dynamic as a database. The issue is about trust, or the lack of trust. Proving the authenticity of a digital object means providing sufficient evidence to convince an adjudicator that the object that has been retrieved is a faithful representation of what is claimed to be the 'original,' or a reliable representation of the object that was relied upon by the originator. Thus the authors of the Uniform Electronic Evidence Act Consultation Paper in Canada indicated there is a need to shift to the integrity of the record-keeping system, and the emphasis needs to be placed on 'system':

> [24] The 'function' of the best evidence rule is to ensure the reliability, that is to say the integrity, of the record to be produced in evidence. It is presumably easier to tell that an original paper record has been altered than to determine any alteration by viewing a copy. In the electronic world, there may or may not be any original paper version of the electronic record. Therefore, the search for integrity of an electronic record has to proceed in another way.
>
> [25] As Ken Chasse said in his 1994 paper for the Conference, at para 46,
>
> > ' … the law should move from "original" to "system", that is, from a dependence upon proof of the integrity of the original business document to a dependence on proof of the integrity of the record-keeping system. This means that the best evidence rule loses most or all of its application in this field …'
>
> [26] Stated another way, the integrity of the record-keeping system is the key to proving the integrity of the record, including any manifestation of the record created, maintained, displayed, reproduced or printed out by a computer system.[4]

1 Geoffrey Yeo, 'Trust and Context in Cyberspace, Archives and Records' (2013) 34 The Journal of the Archives and Records Association 214.

2 For a discussion about the use of terms and meanings, see Sarah Mocas, 'Building theoretical underpinnings for digital forensics research' (2004) 1 Digital Investigation 68.

3 Clifford Lynch, 'Authenticity and integrity in the digital environment: An exploratory analysis of the central role of trust', in *Authenticity in a Digital Environment* (Council on Library Information Resources 2000).

4 Uniform Electronic Evidence Act Consultation Paper (March 1997), available at <www.ulcc.ca/en/1997-whitehorse-yt/377-civil-section-documents/360-electronic-evidence-act-consultation-paper>.

7.85 The Canadian Uniform Electronic Evidence Act offers sound guidance in relation to the evidentiary foundations of electronic evidence, and, in art 6, provides for relevant standards to be considered, although they are not mandatory:

> 6. For the purpose of determining under any rule of law whether an electronic record is admissible, evidence may be presented [in any legal proceeding] in respect of any standard, procedure, usage or practice on how electronic records are to be recorded or stored, having regard to the type of business or endeavour that used, recorded or stored the electronic record and the nature and purpose of the electronic record.

7.86 During the consultation process for the Canadian Uniform Electronic Evidence Act, requests were made to provide for a statutory presumption of reliability based on a standard, but this was rejected on the grounds that it constituted a higher standard for admissibility than was necessary, and the use of a standard might prevent a proper scrutiny of electronic evidence as to weight.[1]

1 MacNeil, *Trusting Records* 53.

7.87 Digital signatures seem to provide a solution to the problem of proving authenticity in the very narrow sense of the integrity of the record and purported authorship of the record. However, they are not used widely. The take-up of digital signatures has been slow, partly due to the problems of authenticating the public key holder.[1] Instead, the use of user names with passwords has become more widespread. If the holder of the user name and password can be identified, and there is no evidence of fraudulent use of the user name, then this can be a good method of authentication, regarding access to the system in which digital objects are stored.

1 This is discussed in detail in Mason, *Electronic Signatures in Law*, and, from the technical (not legal) point of view, Jean-François Blanchette, *Burdens of proof: Cryptographic culture and evidence law in the age of electronic documents* (MIT Press 2012).

7.88 The following section sets out, in general terms, some of the issues that are relevant to the question of the authenticity of electronic evidence.

Technical considerations relating to authenticity

Method of preservation

7.89 Preservation of data is important, especially where electronic evidence could potentially be deleted as part of an ordinary business process.[1] Several methods are used to preserve digital data.[2] Risks attach to whichever method is used, and it is important to ensure that whatever method is employed, it can be defended should the data be the subject of a legal challenge as to its authenticity. Methods include:

> *Technology preservation*: This is the creation of a methodology to conserve the environment in which the data files are set. This includes saving the software and hardware to enable a user to obtain access to and read the data. This is a short-term solution that assumes that people in the future will want to use old hardware and software.
>
> *Technology emulation*: This can take different forms. In essence, this is a method to run the original data and software on a new or current platform. This is achieved by running original software on the new platform that emulates the original platform. Detailed information about the original environment must be stored alongside the digital data itself. Such methods are difficult to develop, and the authenticity of the data will depend on the links between the emulator and

the emulated system. This method also relates to the format in which the data is encoded. The data is copied to the latest form of storage media. Data is converted from one file format (which can no longer be read using the current software) to another format (which is readable with current software). Where the document has been part of a software migration, evidence will be required setting out why migration took place, the methods used to effect the migration, how the quality of the document was validated after migration, and records will be required setting out the names of the people undertaking the exercise, what they did and when they did it.

Data refreshing: This is where data is copied from one set or copy of the digital media to another of the same kind. It can also involve the copying of the data between media of the same type, or to a different kind of media.

1 See Michael H Dore, 'Forced Preservation: Electronic Evidence and the Business Records Hearsay Exception' (2010) 11 Colum Sci & Tech L Rev 76.

2 See also Charles M Dollar, *Authentic Electronic Records: Strategies for Long-Term Access* (Cohasset Associates, Inc, 2002) 29–33 and ch 2; *The State of Digital Preservation: An International Perspective Conference Proceedings* (CLIR pub107, Documentation Abstracts, Inc., Institutes for Information Science, Washington D.C. 24-15 April 2002) 4–31, available at <www.clir.org/pubs/reports/reports/pub107/pub107.pdf>; Oya Y Rieger, *Preservation in the Age of Large-Scale Digitization: A White Paper* (CLIR pub141, February 2008), available at <www.clir.org/pubs/reports/reports/pub141/pub141.pdf>; Jeremy Leighton John, *Digital Forensics and Preservation* (DPC Technology Watch Report 12, 3 November 2012).

Essential technical considerations

7.90 The technical considerations will be a matter of evidence, and often the technical evidence gathered and tendered may simply provide circumstantial evidence. To be authentic, the evidence must establish the identity of the digital document, and must show the integrity of the document.

7.91 *Identity*: The identity of a digital document will need to be established, such as the name of the purported author,[1] the date it was created, the place of origin and the subject matter. It can be argued that this information forms part of the reliability of the document, meaning that if it can be identified correctly, there is a degree of certainty about the document that could be relied upon. Technical evidence may be tendered to provide circumstantial evidence of the identity. For example, if a person says he did not write and send an email, evidence of the fact that he was logged onto a particular computer and at around the same time obtained access to online systems to which only he had access, may go towards showing that that person did, in fact, write the email.

1 Determining the identity of an author is not necessarily a simple process: Gaurav Gupta, Chandan Mazumdar, M S Rao and R B Bhosale, 'Paradigm shift in document related frauds: Characteristics identification for development of a non-destructive automated system for printed documents' (2006) 3 Digital Investigation 43–55; Carole E Chaski, 'Who's at the keyboard? Authorship attribution in digital evidence investigations' (2005) 4 Intl J of Digital Evidence.

7.92 *Integrity*: Integrity is considered to refer to the 'wholeness and soundness' of the document. This in turn is related to whether the document can be considered to be complete and uncorrupted '… in all its essential respects during the course of its existence'. BS ISO 15489 (2001), on the other hand, provides that integrity refers to the record being complete and unaltered. While these definitions of 'integrity' might relate to the ability to verify that the content of a document has not been changed since

it was written, finished and adopted by the author (if the author is known, or remains anonymous for good reasons), it might be necessary to consider other matters, including, but not limited to: whether a time stamp was used, and if so, whether it can be considered to be accurate, and if in doubt, what standards were observed with the particular type of time stamp used; whether it is a partially written document; whether the test for integrity of the document should only apply to the first-in-time version (also called 'first instantiation'[1]), or whether any tracking regarding the document's subsequent circulation is necessary. Following from this, the integrity of the circulation metadata may need to be established, including whether the metadata can be accepted as reliable and meaningful.[2] The concept of integrity will be closely related to the organization's control over the preservation of a document. Underlying the integrity of a document will be the use of digital signatures to provide evidence of verification that the document has not been altered, and the integrity of any digital signatures may also be questioned. It may also be necessary to consider the relevance of any data logs that might exist. Data logs, though complex, have the potential to support or undermine the truth of a claim as to the actions that were being carried out on a particular computer or system at a material time.[3]

1 Teppler, 'Testable Reliability' 255, 220, where the author correctly suggests that 'The data (or information) actually read or perceived by a human reader (or members of a jury) should therefore be considered the last "view" in a set of "views of views" and not the "source" or origination data'.

2 See papers produced by the Recordkeeping Metadata Project, Records Continuum Research Group, Monash University, available at <www.infotech.monash.edu.au/research/groups/rcrg/>; Metadata at <www.ukoln.ac.uk/metadata/>; Dublin Core Metadata Initiative at <http://wiki.dublincore.org/index.php/User_Guide>; Australian Government Recordkeeping Metadata Standard at <www.naa.gov.au/records-management/publications/>; *Model requirements for the Management of Electronic Records, Update and extension* (Office for Official Publications of the European Communities 2008); note relevant articles published by *Ariadne*, available at <www.ariadne.ac.uk>.

3 Karen Kent and Murugiah Souppaya, *Guide to Computer Security Log Management* (National Institute of Standards and Technology, Special Publication 800-92, September 2006) 2.1.3, fourth bullet point, available at <http://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-92.pdf>, although it should be noted that the authors of this text repeat the assumption, which tends to be a problem with many technically literate people, that a data log has the capacity of identifying 'who has used the application and when each person has used it'. All the data are able to demonstrate is the probable use of an application or system by a person who used a password or authentication token. It does not follow that the person using the password or authentication token was the person that was issued with or identified by the password or token that was actually used; see also the discussion in Paul, *Foundations of Digital Evidence* ch 4.

Organizational characteristics

7.93 Procedural controls provide circumstantial evidence of the integrity of a document in digital form. Where policies and procedures are followed, a degree of trust is created that acts to reinforce the probability that a document can be trusted. However, the assumption of integrity cannot be sustained where the procedures are tested in a court and found wanting by the adjudicator.[1] This is why some or all of the following are relevant: the controls put in place to prevent the modification or editing of the record; evidence of the controls to support that the document is authentic by the production of credible metadata, audit trails[2] and relevant reports; the procedures in place to assess and maintain the authenticity of the document over the period of time it has been preserved, including where the document was created, the reason it was created, the technical and procedural framework in which it was created, for whom it

was intended, when and how it was received by the person to whom it was addressed, and how it related to other records linked to the same matter; and whether evidence is available to demonstrate policies were properly created, and that procedures were subsequently adopted and followed to ensure the policies were correctly implemented.[3]

1 *Denco Limited v Joinson* [1992] 1 All ER 413, [1991] IRLR 63, [1991] ICR 172, [1991] 1 WLR 330, EAT. In this case, Wood J observed that the members of the industrial tribunal were 'extremely critical of the security arrangements made by the employers in connection with the use of the computer' ([1991] ICR 172 at 178) – although the observation was made in the context of security, nevertheless the evidence of a sloppy attitude towards something as important as security serves to indicate that other problems may exist that may go to undermine the integrity and thus authenticity of data held on such computer systems.

2 Caroline Allinson, 'Audit trails in evidence – A Queensland case study, Work in Progress' (2002) 1 Journal of Information Law & Technology, online at <www2.warwick.ac.uk/fac/soc/law/elj/jilt/2002_1/allinson/>; Caroline Allinson, 'Audit trails in evidence – Analysis of a Queensland case study' (2003) 2 Journal of Information Law & Technology online at <www2.warwick.ac.uk/fac/soc/law/elj/jilt/2002_1/allinson>.

3 For a discussion about the use of digital signatures in this process, see Stefanie Fischer-Dieskau and Daniel Wilke, 'Electronically signed documents: Legal requirements and measures for their long-term conservation' (2006) 3 Digital Evidence and Electronic Signature Law Review 40; MacNeil, *Trusting Records* 97–104; Thiago A Ramos, Nelson da Silva et al, 'An infrastructure for long-term archiving of authenticated and sensitive electronic documents', in Jan Camenisch and Costas Lambrinoudakis (eds), *Public Key Infrastructures, Services and Applications 7th European Workshop*, Lecture Notes in Computer Science, Vol 6711 (SpringerLink 2011).

7.94 In the US, in the case of *Lorraine v Markel*,[1] there were significant admissibility problems with the evidence, in particular, none of the documents presented were authenticated by affidavit or otherwise. Grimm J looked at the need for authentication and cited the second edition (1997) of *Weinstein's Federal Evidence Manual* as follows:

> The need for authentication and an explanation of the computer's processing will depend on the complexity and novelty of the computer processing. There are many states in the development of computer data where error can be introduced, which can adversely affect the accuracy and reliability of the output. Inaccurate results occur most often because of bad or incomplete data inputting, but can also happen when defective software programs are used or stored-data media become corrupted or damaged.
>
> …
>
> Factors that should be considered in evaluating the reliability of computer-based evidence include the error rate in data inputting, and the security of the systems. The degree of foundation required to authenticate computer-based evidence depends on the quality and completeness of the data input, the complexity of the computer processing, the routineness of the computer operation, and the ability to test and verify results of the computer processing.[2]

1 241 F.R.D. 534 (D. Md. 2007), 73 Fed. R. Evid. Serv. 446, 2007 WL 1300739 (D.Md May 4, 2007), 2007 ILRWeb (P&F) 1805; Brian W Esler, 'Lorraine v Markel: unnecessarily raising the standard for admissibility of electronic evidence' (2007) 4 Digital Evidence and Electronic Signature Law Review 80.

2 241 F.R.D. 534 (D. Md. 2007), 543 and 544.

7.95 Grimm J stated that a witness must 'provide factual specificity about the process by which electronic evidence is created, acquired, maintained, and preserved without alteration or change, or the process by which it is produced if the result of a system or process that does so, as opposed to boilerplate, conclusory statements

that simply parrot the elements of the business records exception, … or public record exception.'[1] The judge referred to the Federal Rules of Evidence, Article IX Authentication and Identification, Rule 901(b), which provides illustrations of how evidence can be authenticated, including authentication or identification (Rule 901(b)(3)) by comparison by the trier of fact or expert witnesses with specimens that have been authenticated and documents, including emails and other electronic records, which can be authenticated or identified by 'appearance, contents, substance, internal patterns, or with circumstances' (Rule 901(b)(4)).

1 241 F.R.D. 534 (D. Md. 2007), 545 and 546.

7.96 Metadata can also be used to authenticate evidence, and the judge noted that:

> Federal Rule of Civil Procedure 34 permits a party to discover electronically sourced information and to identify the form or forms in which it is to be produced. A party therefore can request production of electronically stored information in its 'native format', which includes the metadata for the electronic document. Because metadata shows the date, time and identity of the creator of the electronic record, as well as changes made to it, metadata is a distinctive characteristic of all electronic evidence that can be used to authenticate it under Rule 901(b)(4).[1]

1 241 F.R.D. 534 (D. Md. 2007), 547. The judge went on, at 548, to indicate that metadata is not necessarily as conclusive as he stated.

7.97 Judge Grimm recognized that authenticating electronically stored information presents a myriad of concerns because 'technology changes so rapidly' and is 'often new to many judges'.[1] Further, the 'complexity' or 'novelty' of electronically stored information, with its potential for manipulation, requires greater scrutiny of 'the foundational requirements' than letters or other paper records.[2]

1 241 F.R.D. 534 (D. Md. 2007) at 544.
2 241 F.R.D. 534 (D. Md. 2007), 543–544, quoting Jack B Weinstein & Margaret A Berger, *Weinstein's Federal Evidence* § 900.06[3] (Joseph M McLaughlin ed., 2nd edn., Matthew Bender 1997).

7.98 For a reminder of the importance of the mix of organizational and technical issues that combine to make up for the trustworthiness of evidence in digital form, the Report of the Somalia Commission of Inquiry in Canada offers some stark insights:

> Document disclosure remained incomplete throughout the life of the Inquiry. It took the form of a slow trickle of information rather than an efficient handing over of material. Key documents were missing, altered, and even destroyed. Some came to our attention only by happenstance, such as when they were uncovered by a third-party access to information request. Some key documents were disclosed officially only after their existence was confirmed before the Inquiry by others. Representatives from SILT were reminded continuously of the slow pace and incomplete nature of disclosure. Following numerous meetings on the document transmittal process and private meetings with SILT officials at which we expressed frustration with the process, there were still few results. Finally, faced with altered Somalia-related documents, missing and destroyed field logs, and a missing National Defence Operations Centre computer hard drive, we were compelled to embark on a series of hearings devoted entirely to the issue of disclosure of documents by DND and the Canadian Forces through DND's Directorate General of Public Affairs, as well as to the issue of compliance

with our orders for the production of documents.[1]

1 *Dishonoured Legacy: The Lessons of the Somalia Affair,* Report of the Commission of Inquiry into the Deployment of Canadian Forces to Somalia (Ottawa 1997) Minister of Public Works and Government Services Canada, Executive Summary: Sources of Information, para 3, available at <http://qspace.library.queensu.ca/handle/1974/6881>.

7.99 Of the recommendations published by the Somalia Commission, Recommendation 39.1 required the Department of National Defence to ensure that the National Defence Operations Centre logs are properly maintained in future by implementing the following:

> (a) an audit procedure to ensure that standing operating procedures provide clear and sufficient guidelines on the type of information to be entered and how the information is to be entered;
>
> (b) an adequate data base system, which includes software controls to ensure accurate data entry in each field and appropriate training for operators and users of this system; and
>
> (c) increased system security to an acceptable standard compatible with the objective of national security, including restricting access to authorized persons using only their own accounts and passwords and extending the use of secure (hidden) fields to identify persons entering or deleting data.[1]

1 <http://publications.gc.ca/collections/collection_2015/bcp-pco/CP32-66-1997-eng.pdf>.

7.100 Given the fact that public documents are increasingly being created and stored by bureaucracies across the world in digital form,[1] there is now a degree of pressure upon the keepers of public records to have sufficiently robust systems in place to store documents in digital form. But it is also incumbent on lawyers to be aware of the difficulties surrounding evidence in digital form, because they may need to be more circumspect when agreeing to admit documents that can be adduced into evidence by virtue of their public nature, or because a statute provides for the admissibility of a document that purports to be authentic.[2] One example may serve to illustrate the nature of the problem. Health records are increasingly being put into digital form, and health professionals are increasingly being required to add information to the record electronically. The physical piece of paper is disappearing in the health world, but a serious problem will undoubtedly become manifest in the future, and that relates to the proof of which nurse or doctor entered a particular entry into an electronic record. If an entry is in dispute, proof will be required that a particular person made the entry. This is where the method of authentication will be crucial. Many systems rely on the use of digital signatures stored on smart cards for this purpose, but unless different mechanisms are used to ensure the actual person with the digital signature is the person using the signature as a means of authentication, then it cannot be assumed that the person who purported to make an entry was the actual person who made the entry.[3] This is because many health professionals, for instance, frequently use other professionals' usernames and passwords to log on to systems, as a matter of convenience. Those agreeing, either explicitly or implicitly, to the use of their personal means of authentication by others, fail to understand the significance of what they have agreed to; and if their details are used without their knowledge or permission (for instance, where the practice is generally accepted within an organization or department), then the difficulties in establishing who was responsible for inputting text becomes manifest. Likewise, the National Electronic Conveyancing System in

Australia will see the advent of digital transfers of title and digital certificates of title. The challenge will be to ensure that the security and authentication around transfers of title are adequate to minimize the risk of fraud.[4]

1 Cases in Canada have held that documents purporting to be issued by the Motor Registry can be authenticated, in the absence of evidence to the contrary: *R. v Finnie Distributing (1997) Inc*. [2004] OJ No. 4513, 2004 ONCJ 256.
2 Richard J Matthews, 'When is case law on the web the "official" published source? Criteria, quandaries, and implications for the US and the UK' (2007) 2 Amicus Curiae 19.
3 C R Weir, J F Hurdle, M A Felgar, J M Hoffman, R Both and J R Nebeker, 'Direct Text Entry in Electronic Progress Notes' (2003) 42 Methods of Information in Medicine 61; Simone van Esch, 'The electronic prescription of medication in a Netherlands hospital' (2006) 3 Digital Evidence and Electronic Signature Law Review 55.
4 *National Electronic Conveyancing System*, National E-Conveyancing Development Ltd <www.necs.gov.au>.

Authentication in some special cases

7.101 Where the authenticity of a digital object is in issue, the range of considerations to be taken into account will differ, according to the nature of the evidence to be authenticated and where the evidence is to be found. In the majority of cases, oral and circumstantial evidence will be sufficient to provide for the authenticity of most documents in digital form.[1]

1 Electronic evidence is also capable of corroborating the testimony of a witness, for which see Beryl A Howell and Brian M Heberlig, 'The Lamar Owens case: How digital evidence contributed to an acquittal in an explosive rape case' (2007) 24 The Computer & Internet Lawyer 1; see the discussion regarding the admission of fMRI scans in Neal Feigenson, 'Brain Imaging and Courtroom Evidence: On the Admissibility and Persuasiveness of fMRI', in Michael D A Freeman and Oliver R Goodenough (eds), *Law, Mind and Brain* (Ashgate 2009).

Social networking websites

7.102 Evidence from social networking sites is now regularly admitted into both civil and criminal proceedings. In criminal proceedings, the prosecution has to provide proper foundations for the evidence to be admitted, including the authenticity of the evidence. It might appear that evidence from social networking websites is adduced with what a layman might think is a perfunctory nod towards laying the foundations.[1] Benjamin Greenstone, a history student, claimed that this occurs regularly in the criminal courts in England.[2] The author indicated a wide range of problems that might occur in relation to the evidence to be found on social networking websites, but failed to understand that if the defendant challenges the evidence, then the prosecution will be put to proof before the evidence is admitted. That such evidence is admitted regularly and with little apparent concern for authenticity probably reflects the instructions by the defendants to their lawyers, and also enables the trial to continue without the need for costly authentication exercises that are not necessarily required.[3] This is well understood by lawyers and judges alike, as observed by D. W. Elliott:

> It is not for the opponent to seek its rejection by the judge by showing that it *might* be a forgery: rather should he seek, by cross-examination of the sponsoring witness or by rebutting evidence, to have it rejected by the jury because it *is* a forgery.[4]

1 For a discussion of the position in the US, see Heather L Griffith, 'Understanding and Authenticating Evidence from Social Networking Sites' (2012) 7 Washington Journal of Law, Technology & Arts 209.
2 Benjamin Greenstone, 'Social networking websites as evidence', *Counsel* (April 2012) 25.
3 Micheál O'Floinn and David Ormerod, 'Social networking material as criminal evidence' [2012] Crim LR 486 for a discussion of the approach taken in the courts in England relating to data from social networking sites.
4 D W Elliott, 'Mechanical aids to evidence' [1958] Crim LR 5, 7.

7.103 In another example, entries of a general nature that were made by a complainant on a social networking website in *Regina v D*[1] and that were discovered after the trial were not considered to be relevant to the issues the members of the jury had to determine. In a murder case, the Court of Appeal of Texas held that information from a social networking website controlled by the appellant was sufficiently authenticated, because the information on the relevant web pages was directly linked to the appellant, including his name, address, photographs and references to the murder victim, the arrest of the appellant and comments indicating there was more than one person involved in the gunfire. Morris J commented:

> This type of individualization is significant in authenticating a particular profile page as having been created by the person depicted in it. The more particular and individualized the information, the greater the support for a reasonable juror's finding that the person depicted supplied the information.[2]

1 [2011] EWCA Crim 2305, 2011 WL 4832463.
2 *Tienda v The State of Texas*, 2010 WL 5129722 (Tex.App.-Dallas), 5.

7.104 In *R v Ben-Rejab and Baccar*,[1] the prosecution was permitted to cross-examine the victim on entries she made on a social networking website after the alleged assault to demonstrate that some of the comments she made in her victim impact statement were not true or exaggerated. Michael Ruse, prosecuted for assault causing actual bodily harm, wrote a comment on a social networking site shortly before the members of the jury were due to retire. He wrote 'Another week at court!', and replied to a question from a friend 'Yeah I think I get away with it tbh [to be honest] x,' adding it was 'looking good.' This exchange was printed out and delivered to the court anonymously, where it was handed to the prosecution. Apparently the accused used the name Michael Miles online, and when presented with the evidence, he changed his plea to guilty. In sentencing, Pearson J is reported to have said 'You pleaded guilty part way through the trial only really because you were stupid enough to put on Facebook what amounted to a full confession. Your stupidity really is not much mitigation.'[2]

1 [2012] 1 Cr. App. R. 4, [2011] EWCA Crim 1136.
2 Gareth Bethell, 'Facebook boast to his friends lands dopey thug ABH conviction' (*The News*, 6 June 2012) <www.portsmouth.co.uk/news/local/facebook-boast-to-his-friends-lands-dopey-thug-abh-conviction-1-3919439>.

7.105 Technical evidence can be used to undermine a suggestion – which can be far fetched – that somebody else was responsible for actions made on social networking sites. The case of *The Bussey Law Firm PC v Page*[1] serves to illustrate this point. This was a claim in defamation arising from Internet abuse. The second claimant was Mr Timothy Bussey, a lawyer practising in the state of Colorado in the United States of America. Someone posted a defamatory allegation about him on his Google Maps profile. Jason Page was identified as the person responsible for posting the comments. At issue at trial was whether Mr Bussey could prove to the requisite standard that Mr

Page was responsible for the original posting on 27 January 2012. It was admitted that the posting had been made from Mr Page's Google account, although as Eady J indicated, 'He could hardly do otherwise. This fact was originally established, at no doubt considerable expense, by Mr Bussey who had instructed a firm of California lawyers to obtain a subpoena in respect of Google's records.'[2] Mr Page advanced a number of hypothetical explanations as to how an unidentified third party might have posted the allegations via his account but without his knowledge. The main possibilities were that a third party must have hacked into his Google account in order to post the offending review, and that he or she *might* have been seeking retribution for some decision or action taken by Mr Page in his capacity as moderator of 'sub-reddits' on the www.reddit.com website.[3] Eady J had to concern himself with what the most likely explanation was, on a balance of probabilities. He set out his reasoning:

> 11. ... The likelihood is, in the absence of any convincing explanation to the contrary, that the posting from Mr Page's account was authored or authorised by him. It is extremely improbable that anyone successfully hacked into that account on 27 January 2012 with a view to posting the words complained of. There is no evidence that anyone did so on that date and, moreover, no reason why anyone with a grudge against the Claimants should attempt to go down that route in any event. Why Mr Page should himself choose to attack the Claimants is also unclear, but the most likely explanation would appear to be a purely financial one. I do not need, however, to come to a conclusion on motive since it is not essential to the Claimants' cause of action. All I need say is that the overwhelming probability is that he is responsible for the posting from his account on the date in question and for its remaining accessible thereafter. There is simply no other reasonable explanation.

1 [2015] EWHC 563 (QB).
2 2015] EWHC 563 (QB), [2].
3 2015] EWHC 563 (QB), [5]–[10].

7.106 In this instance, a great deal of relevant technical and related evidence served to demonstrate the impossibility of the claims made by the perpetrator of the defamatory comments.

7.107 There have been circumstances where evidence from social networking sites is admissible, notwithstanding a lack of technical evidence. In the New Brunswick case of *R v Soh*,[2] the Crown tendered images taken of a Facebook conversation, along with photographs, both of which were taken by a police officer from the complainant's computer. The Crown contended that the photographs were real evidence, while the defence maintained the documents were 'electronic documents' containing hearsay, and not admissible. The defence also argued that it could not be proved that the user with the username 'Galuce Soh' was indeed the accused. No evidence was called to explain how Facebook worked or to show that the person with the username was indeed the accused.[2]

1 2014 NBQB 020.
2 It seems that although it is possible to change Facebook Messenger chats, it would require some technical skills: Peter Sayer, 'Attackers could have rewritten logs of their Facebook Messenger chats with you to introduce falsehoods and malicious links' *IT World* (9 June 2016) <www.itworld.com/article/3080859/hackers-could-have-changed-facebook-messenger-chat-logs.html>.

7.108 The judge, LaVigne J, found that the electronic documents system in which the electronic document was recorded or stored was reliable and the electronic system was operating properly at the time, since no evidence to the contrary was presented and there was no other reasonable ground to doubt the integrity of the electronic documents system.[1] It followed that the electronic document print-out satisfied the best evidence rule. The judge was satisfied on the evidence that the document was what it was purported to be, that is, the print-out of a Facebook conversation between the complainant and a person who used the account, for which the username was 'Galuce Soh'. The screen capture print-outs were held to be admissible as electronic documents. The photographs were held to be real evidence, and it was for the members of the jury to give the circumstantial evidence the weight they chose. As to whether the accused was the same person as user with the username 'Galuce Soh', no evidence was called from the provider as to the email address or IP address linking the account to an individual. LaVigne J concluded:

> However, in order to obtain this kind of information, an order requiring the provider to disclose the identity of the person who used a specific IP address to send a message would have to be obtained from a judge every time. I find that this information, though it could prove very useful in identifying a user, is not absolutely necessary in order to prove the user's identity in every case. Furthermore, even if the information was available and linked the account to the accused, the Court would have to be satisfied that the accused himself was using the account at the relevant time.[2]

1 2014 NBQB 020, [30].
2 2014 NBQB 020, [36].

7.109 Based on the evidence, the Crown had established, on a balance of probabilities, that the accused did write and send the messages to the complainant. It was for the jury to decide whether the accused was the author of the messages.[1]

1 2014 NBQB 020, [40].

Email

7.110 Friedman DJ made a similar point to the general observations noted above in the case of *United States of America v Safavian*,[1] where the defendant argued against the trustworthiness of emails that were in turn part of a series of emails.[2] The judge said:

> The Court rejects this as an argument against authentication of the e-mails. The defendant's argument is more appropriately directed to the weight the jury should give the evidence, not to its authenticity. While the defendant is correct that earlier e-mails that are included in a chain – either as ones that have been forwarded or to which another has replied – may be altered, this trait is not specific to e-mail evidence. It can be true of any piece of documentary evidence, such as a letter, a contract or an invoice. Indeed, fraud trials frequently center on altered paper documentation, which, through the use of techniques such as photocopies, white-out, or wholesale forgery, easily can be altered. The *possibility* of alteration does not and cannot be the basis for excluding e-mails as unidentified or unauthenticated as a matter of course, any more than it can be the rationale for excluding paper documents (and copies of those documents). We live in an age of technology and computer use where e-mail communication

> now is a normal and frequent fact for the majority of this nation's population, and is of particular importance in the professional world. The defendant is free to raise this issue with the jury and put on evidence that e-mails are capable of being altered before they are passed on. Absent specific evidence showing alteration, however, the Court will not exclude any embedded e-mails because of the mere possibility that it can be done.[3] (emphasis in the original)

1 435 F.Supp.2d 36 (D.D.C. 2006).
2 In February 2015, a prisoner escaped from jail after sending forged emails to prison authorities that he was on bail: 'Fraudster fools prison officers into thinking he had been granted bail', *The Guardian* (London, 27 February 2015) <www.theguardian.com/uk-news/2015/feb/27/fraudster-neil-moore-escapes-fools-prison-officers>.
3 435 F.Supp.2d 36 (D.D.C. 2006), 41.

7.111 Where the authenticity of data is challenged – for instance, where a party claims he did not write emails – then it might be necessary to consider a forensic analysis of the machine upon which the data is stored. In the case of *Takenaka (UK) Ltd and Corfe v Frankl*,[1] three defamatory emails were sent to Takenaka, ostensibly by a Christina Realtor. At trial, the issue was whether or not David Frankl, a former employee, sent them, or he caused them to be sent. After a number of Norwich Pharmacal proceedings against various Internet Service Providers, Takenaka concluded that the emails originated from a computer located in Turkey belonging to Thames Water, who were engaged a project relating to the Izmit Dam. After leaving his employment with Takenaka, David Frankl secured employment with Thames Water and worked in Turkey on the Izmit project.

1 [2001] EWCA Civ 348.

7.112 It was not in dispute that David Frankl had access to a computer belonging to Thames Water located in Turkey, and that he sent emails on it using the Thames Water account. Thames Water delivered the purported computer to an assistant of a computer expert, Mr Bates, although there was no evidence adduced by Takenaka that it was the computer used by David Frankl in Turkey. In giving his evidence, Mr Frankl claimed that the computer in the hands of the expert was not the one he used, and he described various physical marks that he claimed proved that it was not the computer he used.[1] The trial judge concluded that, notwithstanding this evidence, Mr Frankl used the computer. Mance LJ agreed with this conclusion:

> The reality is that the computer contained ample material making it, in my judgment, effectively impossible to conclude that it was anyone else's computer; if it was somebody else's computer, then one would have thought that the material on it could only have been on it because the contents of the defendant's computer had been copied in their entirety. Be that as it may, it seems to me that there is no plausible or realistic ground for challenging the judge's conclusion on this point shown before us ...[2]

1 [2001] EWCA Civ 348, [7].
2 [2001] EWCA Civ 348, [8].

7.113 After undertaking forensic tests on the computer, Mr Bates concluded that the hard disk had been subjected to extensive corruption on 13 December 1999, a date after the sending of the relevant emails, and while it was in the possession of Thames Water at a time when Mr Frankl no longer had access to the computer. Mr Bates

concluded that it appeared to have been a deliberate attempt to destroy material, and that whoever carried out this activity intended to create the impression of enthusiastic and uninformed searching, browsing and copying. Notwithstanding the damage caused, Mr Bates concluded that the three emails had originated from the computer, and that the author was Mr Frankl. This opinion was based on a number of factors:

> 1. What was left on the hard disk, when considered in combination with the information obtained from the Norwich Pharmacal proceedings.
>
> 2. The timing information relating to the alleged transmission of the emails sent.
>
> 3. The information derived from the computer and the timing of the messages that remained on the hard disk, which indicated that other emails were sent in Mr Frankl's name within a short time of the three emails complained of.

7.114 Mance LJ was fully aware that this was not necessarily conclusive evidence that Mr Frankl was in possession of the computer at the material time, as he indicated:

> In Appendix K at page 78 I note that analysis of computer contents cannot conclusively identify whose finger was on the keyboard at the relevant time. This appendix then goes on to discuss the application of collateral information and present some opinions on the likelihood of more than one operator being responsible for the different types of recovered material. These opinions seem pretty conclusive, but it is essential to consider other ways in which the observed information could have come about. A major presumption here is that anything done to compromise the machine must raise Mr Frankl's suspicions.[1]

1 [2001] EWCA Civ 348, [14].

7.115 In his report, Mr Bates also examined by reference to a number of facts that were known, together with times that would require the user to have physical access to the emails. The known facts included emails sent by Mr Frankl to his wife on the Thames Water account. Part of the report dealing with this issue was cited:

> To make this hypothesis viable
>
> (a) the individual responsible must have been in Turkey at the relevant times.
>
> (b) would have had to have had physical access to the machine at the relevant times.
>
> (c) would need a motive to incriminate Mr Frankl, or
>
> (d) have a grudge against Takenaka (UK) Limited.
>
> (e) would have knowledge of the password in 'davidfrankl' Hotmail account, and finally
>
> (f) have the necessary expertise and foresight to carry out such a convoluted plan.
>
> I do not consider this likely, and I cannot conceive of any alternative hypothesis which would fit the observed facts.[1]

1 [2001] EWCA Civ 348, [16].

7.116 Mance LJ concluded that the evidence demonstrated that the trial judge reached the correct conclusion. Ward LJ concurred.

7.117 Sending emails from someone else's account can be easily done, and for the person affected, it can have serious consequences upon that person's life, not only

because the content of the email sent is potentially damaging, but also due to the fact the person has to take legal action in order to protect his reputation. In the Australian case of *Tassone v Kirkham*,[1] a defamatory email was sent from the plaintiff's work email account to a large number of public servants. Mr Tassone alleged he did not send the email; rather the defendant sent it. The plaintiff alleged that he had left a computer at his workplace logged in to his account. While he was out of the room, the defendant created an email that read: 'hello people, just a note to say that i am a homosexual and i am looking for like minded people to share time with'. The email was sent from Mr Tassone's account and was signed with his electronic signature. The court found that he did not compose or send the email. Supervisors conducted interviews with several work colleagues of the plaintiff, including Mr Kirkham. Mr Tassone took action against Mr Kirkham for defamation. Following the evidence of Mr Kirkham, the court found that he did, on the balance of probabilities, write the email, and found that the content was defamatory, and awarded damages in Mr Tassone's favour.

1 [2014] SADC 134 (7 August 2014).

Email metadata

7.118 There may be cases where the metadata is only part of the evidence before the adjudicator, as in the case of *BSkyb Ltd v HP Enterprise Services UK Ltd (Rev 1)*.[1] In this case, Sky alleged that Joe Galloway, an employee of HP Enterprise Services UK Limited (formerly Electronic Data Systems Limited), created a false email. Ramsey J considered the allegations and all the related evidence, not just the evidence of the email.[2] The judge set out the features of the email that led him to conclude the email was false:

> The 12 July email has a number of features which make it unusual. First, it was not received by Keith Russell or Gary Gordon, according to the email of 26 July 2000 and does not appear to have been received by John Chan. Secondly, it lacks any sign that there was an attachment either in the form of an Icon or text '<>'. Whilst Joe Galloway was correct to state that the way in which attachments are shown varies, there is usually an indication on the face of the email that there was an attachment. Equally on 26 July 2000 John Chan had to add in the attachment because it had obviously not come from the email chain from Joe Galloway. Thirdly, whilst the title to the 12 July email of 'Spreadsheet' would be explicable on the basis that Joe Galloway was copying and pasting an email into the chain the context is strange. It would be expected that Joe Galloway would forward the email sent by him on 12 July at 10:11 and include his apology rather than creating a completely new email with a subject matter 'Spreadsheet'. That is obviously what he did on 26 July because of the use of 'FW: EDS Revised Rate Card 11-7-00.xls' as the subject heading. Equally on 26 July when he 'forwarded' the 12 July email he adopted the unusual method of copying and pasting the email rather than forwarding the email with a subject heading 'FW: Spreadsheet'. Fifthly, if the 'MP Chord Third Pass Revised 12 July.xls' was seen on 20 July 2000 as the relevant commercial rate card then it would have been expected that John Chan, who created it, would have been aware of it and sent it to Barry Yard or that Joe Galloway would have identified that to John Chan as the relevant document or immediately have picked up that John Chan was sending the wrong rate card to Barry Yard when copied into that email.[3]

1 [2010] CILL 2841, 129 Con LR 147, [2010] EWHC 86 (TCC), 26 Const LJ 289, [2010] BLR 267, (2010) 26 Const LJ 289.
2 [2010] EWHC 86 (TCC), [197]–[233].

3 [2010] EWHC 86 (TCC), [232].

7.119 The judge continued:

> Whilst each one of those unusual features might alone not cast sufficient doubt on the genuineness of the 12 July email, taken together they make it implausible that Joe Galloway created and sent the 12 July 2000 email contemporaneously. Having come to the conclusion that I have about his conduct in relation to the Concordia MBA and the evidence that he gave in court, I have no hesitation in finding that Joe Galloway simply created the 12 July email to cover his error in the hope that he could convince everyone that he had spotted the error at the time and dealt with it.[1]

1 [2010] EWHC 86 (TCC), [233]. See *BSkyB Ltd v HP Enterprise Services UK Ltd (No. 2)* 131 Con LR 42, [2010] EWHC 862 (TCC) where Ramsay J reiterated that he found that Joe Galloway made the fraudulent representation and gave perjured evidence, at [42].

7.120 As with all forms of evidence, there are more factors than those available via technical evidence to help determine the authenticity of data in digital form.

Tampering with emails

7.121 Tampering with emails can affect the authenticity of emails, and this can be a complex issue. Consider the New York case of *CAT3, LLC v Black Lineage, Inc.*,[1] which concerned various allegations infringing intellectual property rights. In a motion for sanctions based on violations relating to discovery, the defendants alleged that the plaintiffs CAT3 had deliberately altered emails to gain an advantage in the litigation. Emails were tendered which appeared to show that there was more than one version of the email, and the discrepancies between versions lay in the domain name in the email address of the recipient. After CAT3 provided copies of the emails in question in native format by order of the judge, the defendants then requested the data to be subjected to forensic analysis. The analyst produced a report that claimed the process of deletion and replacement of the domain names was not accidental; rather it was 'the result of intentional human action, and not of an automatic or inadvertent computer process'.[2]

1 2015 WL 5559569 (Memorandum and Order dated 21 September 2015); 164 F.Supp.3d 488 (S.D.N.Y. 2016), 2016 WL 154116 (Motion for sanctions based on discovery violations dated 12 January 2016); 2016 WL 3513703 (S.D.N.Y.) (Memorandum for the reconsideration of 12 January decision, dated 26 January 2016); 2016 WL 1584011 (withdrawal of Motion for Sanctions dated 6 April 2016).
2 164 F.Supp.3d 488 (S.D.N.Y. 2016), 2016 WL 154116, 3.

7.122 The plaintiffs tendered evidence from their chief operating officer, general counsel and director of information technology, all who denied knowledge of anyone materially altering any documents. The plaintiffs explained that the migration of their email server from one provider to another may have caused the domain name change. At a later evidentiary hearing on 1 December 2015, the plaintiffs tendered a report from an expert. Two possibilities were put forward to offer an explanation: that a system can be set up such that the server automatically substitutes a particular address when the email is routed from the client through the server, or the change could occur when email is migrated from one system to another, though the plaintiffs conceded that this is not common. Francis IV J considered that the defendant's expert's conclusions were well supported and by contrast, the plaintiffs' expert's evidence

was less than compelling and was merely speculative, and no evidence was offered to support their alternative theories. The judge concluded that there was clear and compelling evidence that the plaintiffs manipulated the emails to gain an advantage in the litigation, but he acknowledged that the evidence was largely circumstantial. Importantly, the judge found that the fact there were near-duplicate emails showing different addresses cast doubt on the authenticity of both. The judge granted the defendants' motion for sanctions, to the extent that the plaintiffs were precluded from relying on the relevant emails, and ordered they pay the attorneys' fees and costs incurred by the defendants in establishing the spoliation and obtaining relief.

7.123 This was not the end of the matter. CAT3 were subsequently able to make contact with the person who performed the migration of the email system.[1] An explanation was provided in the Memorandum in support of the matter being reconsidered:

> On January 23, 2016, Plaintiffs for the first time since 2013 were able to make contact with Mark Jones ('Jones'), who was the individual that performed email migrations for Plaintiffs' predecessors, in 2013. As detailed in his accompanying Declaration ('Jones Dec.'), Jones knows from his own first-hand knowledge that the reason why certain emails appeared in Plaintiffs' sent-box as having been sent from a 'SlamXHype.com' email address while the identical emails appeared in the recipients' in-boxes as having been sent from an 'Ecko.com' email address (the 'Email Discrepancy') resulted from the inherent functionality of a software tool that he personally used for the Plaintiffs as part of a standard, best practice email migration that pre-dated this lawsuit.[2]

1 2016 WL 3513703 (S.D.N.Y.). The detailed technical reasons were set out in Part II, paragraphs 1–5.
2 At 1(1).

7.124 There does not appear to be a decision on this particular application, but a 'Joint Stipulation of Dismissal with Prejudice and Order' was made by Torres J on 6 April 2016, by which all the claims were dismissed, each party was ordered to pay its own costs, and Black Lineage acknowledged that neither CAT3 nor any of their owners or agents were engaged in any discovery misconduct or wrongdoing, and they accordingly withdrew their motion for sanctions.

More complex data

7.125 Where evidence of a system or the way a computer operated is required, the nature of the evidence will be more demanding. Testing reliability for a mainframe will differ from that of a personal computer: the mainframe cannot be seized or moved, which means there are problems demonstrating that the system is working properly (if this is a necessary pre-condition), ensuring that the nature of the electronic evidence is complete, and determining how the other party can test the evidence; whereas the personal device can be the subject of a seizure order or warrant, and thereafter the quality of the inferences to be drawn from a forensic analysis will depend on the trust in the technology used to take an image of the disk, the techniques undertaken to search the disk, the procedures adopted by the investigator and the substance of any conclusions made by the investigator. The hardware forming a local area network or the Internet cannot be seized, which means other issues must be considered, such as how an investigator ensures that he has searched thoroughly for evidence, while at the

same time demonstrate the reliability of the data. Likewise, the use of cloud computing, where data are stored on a 'virtual' machine on a server, means that the whole server cannot be seized, because the server contains many other 'virtual' machines that contain other users' data, which will be subject to privacy, confidentiality, privilege and relevance, and probably even be subject to the laws of different jurisdictions. Further complications are that a cloud contract may provide that it is the cloud provider who owns the data, leading to questions regarding control and custody.

7.126 Issues that may need to be covered and tested include demonstrating the provenance of the source of the data, how it is authenticated, indicating the process by which the data are acquired, and proving the continuity and reliability of the evidence.[1] In terms of archived data, Luciana Duranti and Corinne Rogers suggest that records are considered trustworthy if they are reliable, accurate, and authentic.[2]

1 Peter Sommer, 'Intrusion detection systems as evidence' [2002] CTLR 67; 'Downloads, logs and captures: Evidence from cyberspace' [2002] CTLR 33; Jean-Marc Dinant, 'The long way from electronic traces to electronic evidence' (2004) 18 International Review of Law Computers & Technology 185; Bertrand Lathoud, 'Formalization of the processing of electronic traces' (2004) 18 International Review of Law Computers & Technology 173; IETF RFC 3227/RFC3227 *Guidelines for Evidence Collection and Archiving (Internet Best Current Practices for the Internet Community)*, available at <www.faqs.org/rfcs/rfc3227.html>.

2 Duranti and Rogers, 'Trust in digital records' 522, 525.

7.127 The European Union has adopted a similar approach in the drafting of art 59 of the Directive on payment services in relation to the evidence in banking disputes:[1]

> Evidence on authentication and execution of payment transactions
>
> 1. Member States shall require that, where a payment service user denies having authorised an executed payment transaction or claims that the payment transaction was not correctly executed, it is for his payment service provider to prove that the payment transaction was authenticated, accurately recorded, entered in the accounts and not affected by a technical breakdown or some other deficiency.
>
> 2. Where a payment service user denies having authorised an executed payment transaction, the use of a payment instrument recorded by the payment service provider shall in itself not necessarily be sufficient to prove either that the payment transaction was authorised by the payer or that the payer acted fraudulently or failed with intent or gross negligence to fulfil one or more of his obligations under Article 56.

1 Directive 2007/64/EC of the European Parliament and of the Council of 13 November 2007 on payment services in the internal market amending Directives 97/7/EC, 2002/65/EC, 2005/60/EC and 2006/48/EC and repealing Directive 97/5/EC (Text with EEA relevance) OJ L319, 5.12.2007, 1–36.

7.128 Additional issues arise when such systems are distributed, as explained by Professor Ladkin when referring to the proposed National En-Route Center commissioned by National Air Traffic Services in 1997:[1]

> Features of the NERC system that make it particularly vulnerable are that it is distributed (parts of the system run on many different computers and must communicate reliably with each other to function correctly); concurrent (it performs many different tasks simultaneously); and real-time (complex operations must be performed in step with unfolding events in the world outside). There is no real hub to such a system—it's more reasonable to think of it as a lot of mutually-communicating tasks going on at once in different physical

locations.[2]

In essence, the requirements for authentication of complex systems can be reduced to the five points noted below:[3]

1. The data (both the content and associated metadata) relied upon in any legal proceedings can be shown to be an accurate representation of the prevailing and existing state of those data at the time relevant to the legal proceedings.

2. If the data have changed, for whatever reason, there is an accurate and reliable method of documenting the changes, including the reasons for any such changes from the moment they were identified (and possibly seized) as potential evidence in legal proceedings.

3. The continuity of the data can be demonstrated between the moment the data were obtained for legal purposes and their submission as an exhibit in legal proceedings.

4. Any techniques that were used to obtain, secure and process the data can be shown to have been appropriate for purpose for which they were applied.

5. The technical and organizational evidence demonstrates that the integrity of the data is trustworthy, and can therefore be considered to be reliable[4] and complete (insofar as the data can be complete), which in turn will depend on the circumstances surrounding the data at the time they were identified as of being potentially relevant in legal proceedings.

1 Memorandum by Professor Peter B Ladkin (ATC 20) submitted to the Select Committee on Environment, Transport and Regional Affairs Fourth Report (ordered by the House of Commons to be printed 27 March 1998), available at <www.publications.parliament.uk/pa/cm199798/cmselect/cmenvtra/360-e/36082.htm>.

2 See also Stephen Castell, *The Appeal Report*, prepared for Central Computer & Telecommunications Agency HM Treasury (Eclipse Publications Limited 1990) 32; Peter G Neumann, *Computer Related Risks* (Addison-Wesley 1995) 96–7; Chris Reed and John Angel, *The Law and Regulation of Information Technology* (6th edn, Oxford University Press 2007) 4.3.2 (note that the authors suggest in 4.3.2.3 that the identity of the sender and recipient of a message should be recorded – it will be of great interest to know how this can be achieved with any certainty; and in 4.3.2.4 the authors suggest that the use of a private key of a digital signature can be produced to 'check the identity of the sender' – it is impossible for a private key to identify the identity of the sender unless the sender accepts they affixed the private key to the message – all the private key of a digital signature proves is that the person that knew the password to the private key affixed the digital signature to the message (or the computer was instructed to remember the password): this is a significantly different proposition, as any digital evidence professional will acknowledge, and a number of commercial entities in the Russian Federation understand this distinction to their cost, for which see Olga I Kudryavtseva, 'The use of electronic digital signatures in banking relationships in the Russian Federation' (2008) 5 Digital Evidence and Electronic Signature Law Review 51; and 'Resolution of the Federal Arbitration Court of Moscow Region of 5 November 2003 N КГ-А 40/8531-03-П' (2008) 5 Digital Evidence and Electronic Signature Law Review 149; Paul, *Foundations of Digital Evidence* ch 8.

3 The proposed tests were introduced in the first edition of this book, and have been altered in the light of further knowledge and constructive criticism in subsequent editions. The tests were further refined in the summer of 2016, because they were included in a draft Convention on Electronic Evidence, which was published in (2016) 13 Digital Evidence and Electronic Signature Law Review S1. The tests are also included in the Convention; the following tests were proposed in Ronald J Rychlak, Joanne Irene Gabrynowicz and Rick Crowsey, 'Legal certification of digital data: The earth resources observation and science data center project' (2007) 33 Journal of Space Law 195, 217: (1) proving the accuracy and reliability of the data; (2) proving the accuracy of the data as it was entered into the computer; (3) showing the reliability and capability of the computer hardware/software; (4) illustrating the process used for the computer graphics; (5) and proving the reliability of the final product.

4 Chasse, 'Electronic records as documentary evidence', at 141: 'Electronic records, however, must be judged by the quality of the electronic record system from which they come'; by implication, as mentioned by Professor Dr. Arzt, the auditing and monitoring procedures and documentation of the hardware and software are included in this resume, for which see Clemens Arzt, 'Use of Satellite Imagery in Legal Proceedings' (1999) 24 Air & Space Law 195, 202; note the comments in relation to the vocabulary of authentication of tangible evidence by Rosemary Pattenden, 'Authenticating "things" in English law: Principles for adducing tangible evidence in common law jury trials' (2008) 12 E & P 277; Gregory P Joseph, *Modern Visual Evidence* (1984–2009) para 7.01[3], who contemplates four primary authentication criteria: completeness of data, complexity of manipulation, routineness of entire operation and verifiability of result; Teppler, 'Testable Reliability' 256: 'Adopting a testable reliability standard would require the inclusion of a means to ascertain not only that digital evidence is "what it purports to be," but that such evidence is what it purports to be as of the time that relevance has been asserted, and that such evidence has remained unchanged since that time'.

7.129 It will not always be necessary to prove each element. The extent of the evidence will be governed by the nature of the data and the complexity of the computer systems from which the data is taken. For instance, in a case alleging that a person drove a vehicle in excess of the speed limit,[1] the defence argued that the prosecution must prove the continuity between the machine taking a photograph and the film being removed and developed.[2] Smith LJ articulated the opinion of the court by indicating that there was no need for evidence of continuity. This was because the film carried its own identification data, and he said:

> There is no possibility that the person developing the film and printing the images could change the information on the data block and therefore no possibility that the wrong photographs will be attributed to the case. Mr Shrimpton's submission that there is a possibility that someone might have retouched the photographs or altered them in some way on a computer does not persuade us. Of course we recognize that such things are possible. But, proof of continuity would not assist in the detection of such actions. If the defence were to raise the issue that there had been tampering with the photographs, no doubt the prosecution could be required to identity those involved in the developing and printing processes so that they could be questioned; but without such an issue being raised, we think that evidence of continuity would be quite pointless.[3]

1 *Griffiths v Director of Public Prosecutions* [2007] RTR 44.
2 *Griffiths v Director of Public Prosecutions* [2007] RTR 44, [19].
3 *Griffiths v Director of Public Prosecutions* [2007] RTR 44, [22].

7.130 For trustworthiness, Neumann offers some remarks that are helpful:[1]

> The term *trustworthiness* implies that something is worthy of being trusted to satisfy its specified requirements. The requirements may specify in detail various system properties such as security, reliability, human safety, and survivability in the presence of a wide range of adversities. Trustworthiness thus implies some sort of assurance measures, and is typically *never* perfect.
>
> Trustworthiness needs to be considered pervasively throughout the system life cycle, through system development, use, operation, maintenance, and evolutionary upgrades. It cannot be easily retrofitted into systems that were not carefully designed and developed. It is dependent on technology and on many other factors—the most important of which ultimately tends to be people. (Italics in the original).

1 Peter G Neumann, 'Risks of untrustworthiness', Twenty-Second Annual Computer Security Applications Conference (ACSAC06) (2006), available at <www.csl.sri.com/users/neumann/acsac06.pdf>.

Business records

7.131 Many jurisdictions now include statutory exceptions to the rule against hearsay, one of which is to allow business records to be admitted as an exception to the rule against hearsay where documents form part of the records of the organization. The rationale behind this exception has its origins in the bankers' books rule[1] where records were entered into log books by bank employees which could be relied upon as a record made at that point in time. This exception to the rule against hearsay is an important one when considering authentication of evidence, because the basis on which the law developed over centuries was that employees would literally enter records on to paper within the binding of a book. The justification was that such records would be more reliable than any witnesses' memory. This is to be contrasted with record keeping today, where records are entered and stored on databases, networks, cloud repositories and so on, where there may be far greater incentives to keep false records, where such false entries may not be capable of being detected. The concept of 'trustworthiness' of those records is called into question and it is arguable that evidence surrounding the computer system in which the records were created and stored may need to be obtained. While not articulating the underlying rationale, Lord Phillips illustrated the assumption in *Horncastle, R. v*:

> Business records are made admissible (by s.117 or, where a machine is involved, s.129) because, in the ordinary way, they are compiled by persons who are disinterested and, in the ordinary course of events, such statements are likely to be accurate; they are therefore admissible as evidence because *prima facie* they are reliable.[2]

1 This rule has its origins in the Bankers' Books Evidence Act 1879 (UK), which provided in s 3: 'Subject to the provisions of this Act, a copy of an entry in a banker's book shall in all legal proceedings be received as prima facie evidence of such entry, and of the matters, transactions, and accounts therein recorded'.

2 *Horncastle, R. v* [2010] 1 Cr App Rep 17, [2010] HRLR 12, [2009] UKSC 14, [2010] 2 AC 373, [2010] 1 Cr App R 17, [2010] UKHRR 1, [2010] 2 WLR 2, [2010] 2 WLR 47, [2010] 2 All ER 359, at [2009] UKSC 14 [35].

7.132 It cannot be right that business records continue to have such an advantage. Teppler is correct when he says: 'By categorizing computer-generated information only as a subset of business records, judges have thus been able to avoid the central issues that are uniquely inherent to the authentication of computer-generated information.'[1]

1 Teppler, 'Testable Reliability', 255, 228, citing the Pennsylvania Supreme Court judgment in *Commonwealth v Klinghoffer*, 564 A.2d 1240 (Pa. 1989), 1241–1243.

7.133 The exclusionary rules of the common law were relaxed by the Bankers' Books Evidence Act 1879. This Act provided that copies of entries in bankers' books – that is, ledgers, day books, cash books, account books and all other books kept in the ordinary business of the bank – are considered as prima facie evidence of the matters recorded,[1] subject to a number of requirements before they can be admitted into evidence. As Professor Tapper remarked, the primary purpose was to prevent the business from being disrupted by the need to produce the original books in court.[2] In 1938, the case of a prosecution at a Metropolitan Police Court was commented upon in the *Journal of Criminal Law*.[3] A bank clerk gave evidence, and produced a photograph of the

document. The representative of the accused did not object to the way the evidence was presented, but the commentator of the case raised a number of issues of relevance, the first of which was that the photograph was secondary evidence of the original, which is correct. The commentator to the case then proceeded to consider the rules by which evidence is admitted under the provisions of the Act. First, the provisions of section 5 were noted. Section 5 provides as follows:

> Verification of copy.
>
> A copy of an entry in a banker's book shall not be received in evidence under this Act unless it be further proved that the copy has been examined with the original entry and is correct.
>
> Such proof shall be given by some person who has examined the copy with the original entry, and may be given either orally or by an affidavit sworn before any commissioner or person authorised to take affidavits.

1 In *Job v Halifax PLC* (2009, unreported), Inglis J accepted print-outs of records cut and pasted from log files as evidence of the matters recorded; the trial has held on 30 April 2009 in Nottingham County Court, and judgment delivered on 4 June 2009. The full transcript of the judgment is available, with a commentary by Alistair Kelman, in (2009) 6 Digital Evidence and Electronic Signature Law Review 235.
2 Tapper, *Computer Law* (4th edn, Longman 1989), 407.
3 'Admissibility of a photograph of a banking account' (1938) 2 Journal of Criminal Law 357.

7.134 It was pointed out that no such evidence was tendered in this case, and it was suggested that the photograph was admitted on the basis that 'the camera cannot lie' – which does not follow. Second, citing the comments by Smith LJ in *Hindson v Ashby*,[1] the bank clerk did not give evidence that he took the photograph that was produced, which meant that the image was no more than hearsay. The commentator distinguished the decision in *R v Tolson*[2] because the purpose of the photograph in *Tolson* was to identify her husband, who was accused of bigamy. In the case of the photograph of the bank account, it was claimed that a witness could not say whether the photograph was correct in every detail of that particular account. A further problem with admitting the photograph arose in the light of the provisions of section 4, which reads:

> Proof that book is a banker's book.
>
> A copy of an entry in a banker's book shall not be received in evidence under this Act unless it be first proved that the book was at the time of the making of the entry one of the ordinary books of the bank, and that the entry was made in the usual and ordinary course of business, and that the book is in the custody or control of the bank.
>
> Such proof may be given by a partner or officer of the bank, and may be given orally or by an affidavit sworn before any commissioner or person authorised to take affidavits.

1 2 Ch (1896) 21.
2 (1864) 4 F & F 103, 176 ER 488.

7.135 The commentator suggested that the photograph could not be admitted unless the photographer was an officer of the bank with the necessary knowledge about the books of the bank. Finally the commentator offered the opinion that there would be no requirement for the photograph to be proved where the bank officer producing it had first checked it against the account to which it related, because the Act does not require the person who made the copy to be called as a witness.

7.136 The technology used by banks has altered considerably during the twentieth century, but this did not prevent judges from providing a wide construction to the statute, as in the (criminal) case of *Barker v Wilson*.[1] The Divisional Court was requested to provide an opinion by way of case stated from North Yorkshire Justices sitting at York. The question was whether the justices reached the correct decision that microfilm was included within the definition of 'bankers' books' in accordance with s 9 of the Act. Caulfield and Bridge LJJ were both of the opinion that this was correct. Caulfield J said:

> The justices came to the conclusion – and they put their conclusions in these terms: that they adopted some robust common sense – that section 9 does include microfilm, which is a modern process of producing banker's records. It is probable that no modern bank in this country now maintains the old-fashioned books which were maintained at the time of the passing of the 1879 Act and possibly maintained for many years after 1879.[2]

1 [1980] 2 All ER 81, [1980] 1 WLR 884.
2 (1980) 70 Cr App R 283, 286.

7.137 Bridge LJ reinforced the point:

> The Bankers' Books Evidence Act 1879 was enacted with the practice of bankers in 1879 in mind. It must be construed in 1980 in relation to the practice of bankers as we now understand it. So construing the definition of "bankers' book" it seems to me that clearly both phrases are apt to include any form of permanent record kept by the bank of transactions relating to the banks' business, made by any of the methods which modern technology makes available, including, in particular, microfilm.[1]

1 (1980) 70 Cr App R 283, 287.

7.138 Professor Tapper commended the flexibility of the judiciary to amend a statutory rule in such circumstances.[1] Section 9 has been amended by various enactments, and the relevant section, s 9(2), now reads as follows:

> (2) Expressions in this Act relating to "bankers' books" include ledgers, day books, cash books, account books and other records used in the ordinary business of the bank, whether those records are in written form or are kept on microfilm, magnetic tape or any other form of mechanical or electronic data retrieval mechanism.

1 Tapper, *Computer Law*, 408. See also the decision in *Victor Chandler International v Customs and Excise Commissioners* [2000] 2 All ER 315, [2000] 1 WLR 1296, CA, in which the Court of Appeal adopted an 'always speaking' construction to a statute, taking into account developments that had taken place since the provision was first enacted, even though it created a criminal offence.

7.139 Other statutory exceptions to the hearsay rule are covered in the standard practitioner texts on the subject.

7.140 When considering evidence tendered under the business records exception, it is necessary to be aware that errors can and do occur – accidentally, deliberately, or because of the failure of the software. For instance, for bank business records to be admitted, it must be proved that the book was, at the time the entry was made, one of the ordinary books of the bank; that the entry was made in the usual and ordinary course of business, and that the book was in the custody or control of the bank. The

rule is based on the presumption that businesses rely on certain records in day-to-day operations which give rise to a certain level of trustworthiness.[1] The exception arose on the basis that employees are under a duty to accurately observe, report and record business facts accurately. There is a belief – that is, there is no proof – that records made by employees are reliable because there is an overriding incentive to keep accurate records, and if digital records are relied upon by business, it follows that such records are apparently adequate for legal proceedings. For the record to be accepted under this exception, it must have been made at or near the time of the event or transaction in issue in the proceedings. It is arguable that in many cases, no one person can give evidence as to the creation, content and reliability of electronic evidence. If a senior manager testifies as to the content of a document stored on a networked computer, the IT administrator may also need to be called to give evidence as to the security and integrity of the system upon which the data are stored. The Alberta case of *R. v Lodoen*[2] illustrates the problem. At issue in a trial within a trial was the admissibility of two pages comprising an Operators License Application. The Director of Motor Vehicles Business Support Services, a division or branch of Alberta Registries, apparently a government department, sought to submit a print-out of a microfilm entry of the original paper application. The original paper application was apparently sent in 1994 to a privately owned and operated registry agency operating under the business name 'Registries Plus Inc.'. The Crown offered no evidence from any employee, past or present, from this agency regarding the document. The Director gave evidence of the contemporary procedure, which comprised sending an original document to an private legal entity called Alberta Registries, who in turn sent the document to a business for the purpose of microfilming the document, which, at the time of the hearing, was a private company called 'Critical Control Solutions'. This company retained the original microfilm copy and the original document was destroyed in due course. This was probably what occurred to the original document in this case. Ogle J said the Crown sought, in essence, to admit 'a hard copy of a microfilm copy of an original microfilm copy of an original hardcopy document' (at [10]). The judge ruled that the evidence could not be admitted under s 30 because (i) the evidence failed to offer a guarantee of trustworthiness normally associated with business records; (ii) it was not possible to determine that the original of the document was a record made in the usual and ordinary course of business, and (iii) the evidence failed to satisfy the requirements of subsection 30(3)(a) relating to the authenticity of the copy. The judge also considered the provisions of s 31 regarding microfilm copies, and rejected the evidence on the basis that (i) the evidence was not a business record of Alberta Registries, but a document prepared by a private company in 1994; (ii) it was not possible to determine that the document photographed was the original of the Operator's Licence Application and therefore a business record of the private agency that apparently created it – in such ac case it might have been admissible for that reason, and (iii) the Crown failed to demonstrate compliance with the additional requirements of s 30(1)(e) regarding which agency was responsible for destroying the original document. Finally, Ogle J also determined that the evidence was not admissible under common law because there was no evidence to determine the reliability of the original document.

1 *R v Lemay* (2004) 227 W.A.C. 279.

2 CarswellAlta 1536, 2009 ABPC 274, [2009] A.W.L.D. 4271, [2009] A.W.L.D. 4272, [2009] A.W.L.D. 4273, 14 Alta. L.R. (5th) 130, 480 A.R. 327, 86 W.C.B. (2d) 753.

7.141 The New York case of *Lobiondo v Leitman*[1] illustrates the day-to-day problems that undermine the belief that employees are reliable because there is an overriding incentive to keep accurate records. This was a hearing before Kerrigan J on a motion in limine by the plaintiff for an order admitting certain records into evidence and excluding others. The case concerned medical malpractice. The trial commenced on 26 February 2007. Medical records from Elite Physical Therapy, Inc., a medical service provider that performed physical therapy on Andrea Lobiondo in 2003, were delivered to the court pursuant to a subpoena. The plaintiff's attorney noticed that the evaluation charts and narratives that were printed out from the software and provided in 2007 were different from the ones in the records that the plaintiff had obtained in 2003. The importance of this is set out in the decision by Kerrigan J:

> Plaintiff's counsel had assumed that the subpoenaed records were the same as the records he had obtained in 2003. It is his opinion that the time-line presented by his version of the records supports plaintiff's claim that her left shoulder was injured during the biopsy on March 24, 2003, since it was documented that on March 12 she had complained only of right shoulder pain, but that on March 31, one week after her biopsy, she presented with pain to both shoulders. In contrast, the version furnished to the Court pursuant to subpoena indicates that plaintiff had a left shoulder problem on March 12, 2003, before the biopsy.[2]

1 New York Supreme Court, Queens County, Index number 10037.04, dated 25 September 2007, at <https://nycourts.gov/library/queens/decisions_07-08.shtml>.
2 Page 2.

7.142 The version of the records subpoenaed for the trial was damaging to the plaintiff's case. However, the evidence was such that if she continued with the case, she would have to challenge her own therapist's records. On application by the plaintiff, a mistrial was granted. Leave was granted to conduct a disposition of Elite to ascertain the reason for the apparent discrepancies among the three versions of its records. The purpose was to determine which of the three sets of records should be admitted into evidence at the new trial. The deposition occurred on 8 May 2007. At this deposition, the owner of Elite, Fotis Tsolis, could not explain the discrepancies, other than to suggest there might have been a computer 'glitch'. The following information was ascertained as a result of the deposition:

> (i) How the records were generated: Tsolis jotted down notes and entered the information immediately, but this depended on the time constraints; he would put them into the computer except for prescriptions, which a secretary would input. It was established that there were three versions of the records, and in response, Tsolis said that it was most likely the secretary, although he did not know which secretary would have done this, and did not know what secretaries were working for him in 2003. He also did not know who generated the version of the records pursuant to the subpoena.
>
> (ii) Problems with the computer system: the computers crashed two or three times from 2003 to 2007; on one occasion, a technician from Dell erased whatever was on the computer because of a virus; updates to the computer system had new formats and templates; he had changed some of the wording in patients' reports; there was no explanation as to why there are three separate and distinct sets of computer records.
>
> Records maintained in the regular course of business: Counsel for Andrea Lobiondo failed to ascertain whether the computer records obtained in 2003

> were maintained in the regular course of Elite's business, and Tsolis speculated that a student who had worked for him as an aide or helper may have inputted the information when asked if he could explain why the court's set of records was different from the plaintiff's and defendant's sets.

7.143 The judge summed up the position:

> Tsolis' inability to account for the contradictory and materially different versions of the same purported records, other than to surmise that the existence of three sets containing markedly different information may be attributable to a 'computer glitch' caused by a virus, his inability to identify who actually inputted the information into the computer, his tacit admission that the information may have been inaccurately inputted by an inexperienced student who had worked for him cleaning rooms or doing odd jobs, and that the information may have been changed subsequently to its original entry, leads this Court to conclude that none of the computer records of any of the three versions are reliable as business records and, therefore, none of the computer records of Elite, either those furnished to the Court by subpoena, those obtained by plaintiff's counsel in 2003 or those obtained by defendant's counsel in 2005 are admissible.[1]

The motion in limine was granted in respect of the records in dispute, and the attorneys for the parties were directed to appear for jury selection on 10 October 2007 at 9:30 am.

1 Page 4.

7.144 The problem with data in digital form is that it cannot always be trusted, as noted in the chapter on reliability. For instance, invoices created automatically by software code can be inaccurate;[1] spreadsheet programs are used in organizations or every description every day, yet software errors are prevalent in such programs;[2] affecting the Security Services, where a formatting fault on a spreadsheet program altered the last three digits of each of a number of telephone numbers to '000';[3] the banking sector, which uses spreadsheet programs extensively, has many known highly significant problems;[4] faulty spreadsheet programs are also used in the medical sphere,[5] and mobile telephone records are also a significant area for concern.[6]

1 The inaccuracy of invoices is a significant cause of complaints to the Ombudsman Service Limited, as indicated in an announcement on 8 June 2015: 'Amongst the complaints made between April 2014 to March 2015, consumers were most irritated by billing, which accounted for 85% of all complaints made, these relate to various billing issues such as disputed charges, inaccurate invoices, or consumers not receiving a bill at all'. For which see <www.ombudsman-services.org/ombudsman-services-publishes-new-data-on-energy-complaints-by-supplier.html>.

2 *Review of Medicare Bad Debt and Pitt County Memorial Hospital for the Fiscal Year Ended September 30, 1999* (Department of Health and Human Services, Office of Inspector General, January 2003, A-04-02-02016), available at <https://oig.hhs.gov/oas/reports/region4/40202016.pdf>; see also Stephen G Powell, Kenneth R Baker and Barry Lawson, 'A critical review of the literature on spreadsheet errors' (2008) 46 Decision Support Systems 128; Stephen G Powell, Kenneth R Baker and Barry Lawson, 'Errors in Operational Spreadsheets' (2009) 21 Journal of Organizational and End User Computing 24; Stephen G Powell, Kenneth R Baker and Barry Lawson, 'Impact of errors in operational spreadsheets' (2009) 47 Decision Support Systems 126; Raymond R Panko, 'What We Don't Know About Spreadsheet Errors Today: The Facts, Why We Don't Believe Them, and What We Need to Do', presented at EuSpRIG 2015 (London, 9 July 2015), available at <www.eusprig.org/presentations/Presented%20EuSpRIG%202015%20What%20We%20Don't%20Know%20About%20Spreadsheet%20Errors.pdf>.

3 2010 Annual Report of the Interception of Communications Commissioner (Ordered by the House of Commons to be printed 30 June 2011; Laid before the Scottish Parliament by the Scottish Ministers June 2011), para 7.33.

4 Victoria Lemieux, 'Archiving: The Overlooked Spreadsheet Risk', Proceedings of the European Spreadsheet Risks Interest Group (EuSpRIG) 2005; Grenville J Croll, 'Spreadsheets and the Financial Collapse', Proceedings of the European Spreadsheet Risks Interest Group (EuSpRIG) 2009, both available at <www.eusprig.org/conference-abstracts.htm>.
5 Grenville J Croll and Raymond J Butler, S'preadsheets in Clinical Medicine – A Public Health Warning', available at <www.eusprig.org/conference-abstracts.htm>.
6 See the chapter on software code as the witness.

7.145 A judge will generally accept the authenticity of digital data where evidence is proffered that the system that produced the data is used on a regular basis, and it is also relied upon in the normal course of business, thus confirming that an assertion of reliance (which it usually is) is sufficient to establish authenticity, and therefore the trustworthiness of the data. However, as with provisions permitting business records to be adduced into legal proceedings, Professor Seng has made some pertinent observations in the context of the legislation in Singapore that affects the trustworthiness and reliability, and thus the authenticity, of entries in bankers' books and business records in digital form:

> It is argued that business records are reliable because the statement maker would use such records for the management of his business, and they are necessary because there are no records otherwise kept of business activities. However, it is submitted that evidence of authentication highlights some additional and unique problems with computer output that are not necessarily taken into consideration when the business records exception is applied. If one considers the six steps set out above, the business records exception clearly cannot manage issues such as whether the recorded information has been improperly manipulated or altered, whether the computer is operating properly when recording, storing and extracting such information, and under what circumstances is the information extracted from the computer (steps (iv) to (vi) respectively).
>
> …
>
> The problem with manipulation or alteration of computer records is that unlike documentary records, such manipulations or alterations are almost invariably untraceable. This is inherent in the nature of electronic records. While the business records exception does not preclude the opponent from challenging that the output has been manipulated or altered, because he does not have easy access to the proponent's computer system nor does he have knowledge of it usage policy. Having only the business records exception in effect casts the evidential burden on the opponent to challenge the authenticity of electronic records. This is really an unfair burden, because the opponent is at an information disadvantage …[1]

1 Seng, 'Computer output as evidence' 171–2.

7.146 In *Zezev and Yarimaka v Governor of HM Prison Brixton*,[1] there was an application for *habeas corpus* by the applicants in an attempt to resist extradition to the US. The proceedings were based on six charges, one of which was that of conspiring to cause an unauthorized modification of computer material. Helen Malcolm QC (as she now is) argued (correctly, it is suggested) that there was no evidence to support this particular charge faced by Zezev. The actual basis of the facts relied upon is not clear: the head note suggests that 'There was evidence against the first applicant that he would use the computer so as to record the arrival of information which did not come from the purported source', and the U.S. Attorney's Office for the Southern District of New York indicated that it was an email: 'The evidence demonstrated that ZEZEV sent an email

on April 17, 2000, to Michael Bloomberg threatening that if Michael Bloomberg did not send him $200,000 he would disclose to the media and customers of Bloomberg LP that he was able to gain unauthorized access to Bloomberg's computer system'.[2] For the evidence to be relevant to the charge, it was necessary for the evidence to demonstrate that the threatened act would 'impair the operation of any such program or the reliability of any such data' under s 3(2)(c) of the Computer Misuse Act 1990. In addressing this point, Lord Woolf CJ considered that if an email was sent by Zezev to the Bloomberg server (Zezev used an email account at a company called Hotmail, which Zezev had registered under a false name (the email was traced to Kazkommerts Securities, where Zezev worked)), the receipt of such an email might affect the reliability of the data on the computer, thus ensuring the evidence therefore fell under the provisions of s 3(2)(c). If the proposition was that Zezev was going to send an email from a hidden address to the press and customers of Bloomberg, then it was not correct to say that the reliability of a computer or computer-like device (or the authenticity of the email) was necessarily affected by the fact that such an email was sent and received. In this instance, the contents of any email sent by Zezev would have been accurate. The only aspect of the email that would have been questionable, had it been sent, was the actual source of the email – but even the source of the email (whether from a hidden address or from within the Bloomberg server) would not affect reliability – it would have to be determined what, exactly, would be deemed to be unreliable, and why it was unreliable. If the comments by Wright J indicate that Zezev intended to bypass the relevant security in place to plant an email into the Bloomberg server, then such an action would indeed not be authorized, but it would not affect the reliability of any data.[3] No attempt was made to determine what data would be affected. The editors of *Archbold*[4] refer to the rationale for the decision: 'If a computer is caused to record information which shows that it came from one person, when it in fact came from someone else, that manifestly affects its reliability, and thus the reliability of the data in the computer is impaired within the meaning of section 3(2)(c).'[5] However, the editors do not indicate that there must be a cause and effect – it does not follow that where one person causes information to be recorded that purports to come from another person, that reliability is 'manifestly' affected.[6]

1 [2002] EWHC Admin 589, [2002] 2 Cr App R 33.
2 Press release dated 1 July 2003 that provided a brief outline of the evidence at the trial, available at <www.justice.gov/archive/criminal/cybercrime/press-releases/2003/zezevSent.htm>.
3 [2002] 2 Cr App R 33, [22].
4 (Sweet & Maxwell 2016).
5 At 23–89.
6 For a discussion of the position in the US, see Paul, *Foundations of Digital Evidence*, Part III.

7.147 In summary, in the context of business records, the existing rules of evidence in most jurisdictions allow evidence to be tendered without a threshold test as to its reliability. In the digital age, this is a dangerous position. It is, however, incumbent upon the opposing party to call into question the reliability of the data, in which case it must be proved that the data is what it purports to be. If, as many commentators suggest, the real test of authenticity should be around the integrity of the system in which the data are created and stored, then guidelines and standards that can assist in providing security and integrity can be developed. However, first, the rules of evidence need to recognize this requirement.

Evidence in criminal proceedings

7.148 The previous requirements relating to the authentication of computer evidence have been removed in England & Wales, and the current position is now governed by s 133 of the Criminal Justice Act 2003, which provides as follows:

> 133 Proof of statements in documents
>
> Where a statement in a document is admissible as evidence in criminal proceedings, the statement may be proved by producing either-
>
> (a) the document, or
>
> (b) (whether or not the document exists) a copy of the document or of the material part of it, authenticated in whatever way the court may approve.

7.149 The Explanatory Notes to the Act states, that section 133 'corresponds to the position under section 27 of the Criminal Justice Act 1988, whereby a statement in a document can be proved by producing either the original document or an authenticated copy' and continues, 'It is intended to cover all forms of copying including the use of imaging technology.'[1] Interestingly, the document must be an original or an authentic copy, which illustrates the need to pay careful attention to the means by which a document in digital form is authenticated before the court.[2] The use of imaging technology is also a mechanism of obtaining a copy of the original data, although the actual technology that is used to obtain an image of data may be challenged. The number of removes a copy may be from the original is dealt with indirectly by reference to the meaning of 'copy', which 'in relation to a document, means anything on to which information recorded in the document has been copied, by whatever means and whether directly or indirectly.'[3] This requires the trial judge to determine how a digital document is authenticated, which is why guidance on the mechanisms by which authenticity is tested can be so important. In essence, the move has been towards assessing the weight to be given to electronic evidence.

1 At paragraph 436.
2 O'Floinn and Ormerod, 'Social networking material as criminal evidence'.
3 Section 134(1) Interpretation of Chapter 2.

7.150 The Court of Appeal's stance in *R v Damien O'Connor*[1] has wider implications on the admissibility of electronic evidence beyond its own facts. The appellant and several others were accused of conspiring to import heroin and cocaine into the UK from Belgium. O'Connor was living in Belgium at the time. The prosecution relied upon telephone records provided by the Belgian police in relation to a mobile telephone used by the leader of the conspiracy, but there was no accompanying statement from the Belgian telephone provider. The court concluded that it was arguable that the records, which were produced by the Belgian authorities and handed to the prosecution, were not in fact statements made by a person. Hooper LJ went on to say that if the court was wrong on this point, 'and one concentrates on the person who interrogated the Belgium provider computer and obtained the data for the 8136 phone, and if one assumes that in that respect a person is making a representation for the purposes of section 115, then the issue has to be whether it is admissible under section 117. The judge held that it was.'[2]

1 [2010] EWCA Crim 2287, Times, July 19, 2010.
2 [2010] EWCA Crim 2287, [16].

Concluding comments

7.151 We live in the age of software code – and we have moved well beyond the time when electronic evidence was only extracted from a single physical device, whether a stand-alone computer or telephone – evidence in electronic form transcends individual devices, and the investigative approach needs to reflect this reality, as should the approach that lawyers and judges take to authentication. This topic was discussed during the course of a judicial panel discussion entitled *E-Discovery: Where We've Been, Where We Are, Where We're Going*, hosted by the Ave Maria School of Law on 21 January 2013. The members of the panel included the Hon. Andrew J. Peck, Magistrate Judge for the Southern District of New York, serving as Chief Magistrate Judge in 2004–5; the Hon. John M. Facciola, Magistrate Judge in the District of Columbia, and Professor Steven W. Teppler.[1] Judge Facciola commented that there were two schools of thought: the first, represented by judges who take the simplistic view that 'Authenticity has never been a very difficult hurdle to overcome, so it should not be here' and the second school of thought represented by George L. Paul, who argues that this approach 'does not make any sense, philosophically or any other way'.[2] The judge went on:[3]

> So the case law shows that in terms of authentication, we are still using the traditional model of looking for additional circumstantial evidence that permits a reasonable person to find that the person who is in issue did, in fact, post it. The burden to prove authenticity is very light indeed, and we all know that once the judge determines that a reasonable person could make that determination, the determination of whether he did or not is for the jury. It goes to weight, not to relevance.

1 Steven W Teppler, 'E-Discovery: Where We've Been, Where We Are, Where We're Going' (2014) 12 Ave Maria L Rev 1.
2 Teppler, 'E-Discovery' 53.
3 Teppler, 'E-Discovery' 54.

7.152 Professor Teppler responded:[1]

> There will be truthful information that will be actual information, and there will be very nearly true information (maybe off by a couple of zeroes in the contract) which will go to the weight. How would you like to have a judicial determination based on a gamble? A jury may find, correctly, that one is okay and the other one is not, when it is true, but how do you know that an attorney will be able to competently argue that?

1 Teppler, 'E-Discovery' 54–5.

7.153 That the approach to the authentication of electronic evidence is badly in need of revision may be illustrated by the improper prosecution of nurses in the Princess of Wales Hospital in Bridgend, Wales.[1] It cannot be right that innocent people are investigated and prosecuted because of the failure of manufacturers of IT systems to produce effective systems, and because administrators to fail to understand the weakness of systems they buy and require employees to use – whilst such systems, and thus business records – are open to being manipulated by anybody with access to the record. Andrew Bridgen, Conservative Member of Parliament for North West Leicestershire, also illustrated this in the House of Commons during a debate on 17 December 2014 regarding the Post Office Mediation Scheme and the Horizon system used in all Post Offices. Mr Bridgen repeated an account of a constituent, Mr Michael

Rudkin, who was a sub-postmaster for 15 years when he was invited to a meeting at the Fujitsu/Post Office Ltd offices in Bracknell to discuss problems with the Horizon system:[2]

> On arrival that morning, my constituent signed the visitors' book in reception and waited for his chaperone, a Mr Martin Rolfe. … Mr Rolfe asked Mr Rudkin to follow him through a number of pass card-protected security doors to some stairs. They went down to the ground floor and then entered the boiler room. Mr Rudkin states that a number of men dressed in casual office wear were standing around the doorway. They became very uncomfortable about Mr Rudkin's presence and left.
>
> Having entered the boiler room, Mr Rudkin instantly recognised two Horizon terminals. There were data on both screens, and an operative was sitting in front of one of them, on which the pure feed for the Horizon system came into the building. Mr Rudkin asked if what he could see were real-time data available on the system. Mr Rolfe said, 'Yes. I can actually alter a bureau de change figure to demonstrate that this is live'—he was going to alter a figure in a sub-postmaster's account. He then laughed and said, 'I'll have to put it back. Otherwise, the sub-postmaster's account will be short tonight'. Mr Rudkin expressed deep concern, because he had been told that no one had remote access to a sub-postmaster's account. At that point, he was politely but speedily taken to reception, and he was told to leave the building.[3]

1 Harold Thimbleby, 'Cybersecurity problems in a typical hospital (and probably in all of them)', (forthcoming 2017) Safety-Critical Systems Club; 'Nurses cleared of wilful neglect at Princess of Wales Hospital in Bridgend' (*South Wales Evening Post*, 14 October 2015) <www.southwales-eveningpost.co.uk/nurses-cleared-wilful-neglect-princess-wales/story-27983645-detail/story.html>; 'Princess of Wales Hospital nurse neglect trial collapses' (*BBC News*, 14 October 2015) <www.bbc.co.uk/news/uk-wales-south-east-wales-34527845>. Please see para 9.90 and following in Chapter 9.

2 For more detail, see *Justice for Subpostmasters Alliance* <www.jfsa.org.uk> and <http://becarefulwhatyouwishfornickwallis.blogspot.co.uk/2013/08/post-office-2nd-sight-report-into.html>.

3 Columns 535WH and 536WH; the observations noted about the alteration of data were also made by Richard Roll, a previous employee of Fujitsu in a BBC 1 Panorama programme entitled 'Trouble at the Post Office' broadcast on Monday 17 August 2015 at 7:30 pm – for excerpts from the programme, see <https://ukcampaign4change.com/2015/08/18/post-office-horizon-it-and-last-nights-panorama/>.

7.154 This means that investigative authorities must also accept the reality surrounding the requirement to prove the authenticity and integrity of electronic evidence. This was a consideration addressed by Hallett LJ in the case of *R v Saward*,[1] where the prosecution had to demonstrate the authenticity of intercepted audio evidence that was recorded on a mainframe computer in the Netherlands and then recorded on to a CD. Hallett LJ commented that the court had

> … reservations about the profitability of the four day exercise of putting the Crown to strict proof of the exhibit. All those involved in the conduct of criminal trials must be aware by now of the constraints upon resources as we are far from persuaded that this was proper use of limited resources.[2]

1 [2005] EWCA Crim 3183.

2 [2005] EWCA Crim 3183, [44].

7.155 The problem with this observation is that criminal proceedings demand that the evidence be beyond reproach. We live in a world of technology that is vastly complex,

and it follows that the costs of prosecuting will be far more significant than believed hitherto, and 'cost-effective' administration should not override justice.[1]

1 Stephen Mason and Nicholas Bohm, 'Banking and Fraud', a written submission to the Treasury Committee on 17 January 2011, available at <www.publications.parliament.uk/pa/cm201011/cmselect/cmtreasy/430/430vw25.htm>; David M Paciocco, 'Proof and Progress: Coping with the Law of Evidence in a Technological Age' (2013) 11 Canadian Journal of Law and Technology 181, 190.

7.156 For civil matters, a more robust consideration of authentication is required, but not so stringent that it makes it difficult for authentic evidence to be admitted. The current tests used are based on rules built up over centuries with paper evidence. Those tests need revising in light of the fact that electronic evidence is fundamentally different from paper. Documents that fall within exceptions to the rule against hearsay, such as the business records exception, must not be admitted without a witness applying his mind as to whether the electronic documents are, in fact, authentic or whether there is a risk that they have been altered between the time they were created and the time they are to be admitted into evidence.

8

Encrypted data

Stephen Mason and Alisdair Gillespie

8.1 In any discussions of criminal investigations in the digital era, one cannot avoid the issues of the widespread use of encryption and how this has affected criminal investigations.[1] For a statement on the effect that a criminal act has on the accused, the observations by the Lord Justice-General in the Court of Judiciary in Scotland in the case of *M'Garry v Byrne* remain relevant, and help to put the dilemma of the investigating authorities into perspective when dealing with encrypted data:

> Every man is entitled to the enjoyment of personal liberty, but he forfeits that right by committing crime; and, where the criminal law warrants his arrest on a criminal charge, his personal liberty is unavoidably invaded, not merely by subjecting him to detention, but also to the extent necessary to enable the police to observe and collect the real evidence (afforded by his person, his apparel, or the contents of his pockets) of his connexion with the crime and his identity with the criminal.[2]

1 Dorothy E Denning and William E Baugh Jr., 'Hiding crimes in cyberspace' (1999) 2 Information, Communication and Society 251 (this is an interesting article setting out a number of early cases from across the world in which criminals and terrorists used encryption); Eoghan Casey, 'Practical approaches to recovering encrypted digital evidence' (2002) 1 Intl J of Digital Evidence <www.utica.edu/academic/institutes/ecii/publications/articles/A04AF2FB-BD97-C28C-7F9F4349043FD3A9.pdf>; Jason Siegfried and others, 'Examining the encryption threat' (2004) 2 Intl J of Digital Evidence <www.utica.edu/academic/institutes/ecii/publications/articles/A0B0C4A4-9660-B26E-12521C098684EF12.pdf>; Hank Wolfe, 'Penetrating encrypted evidence' (2004) 1 Digital Investigation 102; Christopher James Hargreaves and Howard Chivers, 'Recovery of encryption keys from memory using a linear scan' in *Proceedings of the 2008 Third International Conference on Availability, Reliability and Security* (IEEE Computer Society 2008) 1369–1376; Carsten Maartmann-Moe and others, 'The persistence of memory: Forensic identification and extraction of cryptographic keys' (2009) 6(Supp) Digital Investigation S132–S140.

1 1933 JC 72, 78.

8.2 It is with this underlying rationale in mind that we explore the topic of encrypted materials in this chapter.

8.3 Encryption (or enciphering) is the process by which a plaintext (or cleartext) message is disguised sufficiently to hide the substance of the content. As well as ordinary text, a plaintext message can be a stream of binary digits, a text file, a bitmap, a recording of sound in digital form, audio images of a video or film, or any other information formed into digital bits. When a message has been encrypted, it is known as ciphertext. The opposite procedure – that of turning the ciphertext back into plaintext – is called decryption (or deciphering). An encryption scheme usually uses a 'key' to encrypt and decrypt the message. Data that is encrypted properly can be virtually impossible to decrypt.

Stephen Mason and Alisdair Gillespie, 'Encrypted data', in Stephen Mason and Daniel Seng (eds.), *Electronic Evidence* (4th edn, University of London 2017) 261–84.

Methods to obtain encrypted data

8.4 Section s 56(1), Regulation of the Investigatory Powers Act 2000 (RIPA 2000) defines a 'key' as follows:

> [a key] in relation to any electronic data, means any key, code, password, algorithm or other data the use of which (with or without other keys) –
>
> (a) allows access to the electronic data, or
>
> (b) facilitates the putting of data into an intelligible form;

8.5 As defined in RIPA 2000, a key includes a password that is used to encrypt the actual key, which either on its own or with other keys, is used to encrypt the electronic data in question.[1]

1 For an example of what a private key looks like, see Stephen Mason, *Electronic Signatures in Law* (4th edn, University of London 2016) 153.

8.6 There are a number of methods that can be used to obtain the key or the password to enable the key to be used to reveal the plaintext:

> (i) The person or under investigation could provide the password or key. In this instance, the key is usually provided in electronic form.
>
> (ii) The investigating authorities could:
>
> (a) Try to 'break' the key or the password with decryption tools. An example is the dictionary attack. Where the software used for encryption or password protection allows multiple entries, it is possible to use a dictionary of words and phrases together with an automated tool to automatically pass each entry in the dictionary in turn to the protected file and see if it succeeds in opening the file. If the required passphrase is in the dictionary, the file will be rendered accessible.
>
> (b) Attempt a 'brute force attack' using powerful computers to try all possible keys[1] or the password.[2] This can occur where the software used for encryption or password protection allows multiple attempts, it is possible to pass it an increasingly complex sequence of characters. Short passphrases using combinations of just a-z and A-Z can often be identified in a reasonable time on low cost equipment. Longer or more complex passphrases may never be correctly 'guessed' even using significant processing power.
>
> (c) A vulnerability attack – this is where the implementation of the encryption or password protection used is flawed and susceptible to programmatic compromise. For example, where the encryption or password protection software compares the entered passphrase with a copy of the correct passphrase that is stored in an insecure way and which can be readily identified by monitoring the use of RAM or disk access.[3]
>
> (d) Use intelligence about an individual to work out the password – the suspect might have used an easy-to-guess password based on a name, a number or a date familiar to him.[4]
>
> (e) Use covertly-installed keylogging software to record the suspect entering the password into the computer to retrieve the password, and then the key.[5]

1 In *R v ADJ* [2005] VSCA 102 the defendant claimed that he could not recall the password, and suggested possible alternatives, none of which were correct, so the police used password cracking software that took over four months to identify the password: the encrypted partition revealed a large quantity of abusive images of children.
2 In *Rollo (William) v HM Advocate* 1997 JC 23, 1997 SLT 958 (HCJ) the police succeeded in gaining access to an encrypted part of a Memomaster notebook by trying a number of combinations, one of which – the appellant's date of birth – was successful; see also *U.S. v Kim* 677 F.Supp.2d 930 (S.D.Tex 2009).
3 We owe the detailed discussion on (a) – (c) to Hein Dries.
4 See Ian Grigg and Peter Gutmann, 'The curse of cryptography numerology' (2011) 9 IEEE Security & Privacy 70 for a brief foray into the failure of everything but the cryptography.
5 *U.S. v Scarfo* 180 F.Supp.2d 572 (D.N.J. 2001). Whilst this may not tell you who depressed the keys (thus proving who had control) it would provide access to the encrypted material which, by itself, is likely to assist the wider investigation. See also Giuseppe Vaciago and David Silva Ramalho, 'Online searches and online surveillance: the use of trojans and other types of malware as means of obtaining evidence in criminal proceedings' (2016) 13 Digital Evidence and Electronic Signature Law Review 13, 88.

8.7 As a final option (in the case of the UK), it is possible to make a request to the National Technical Assistance Centre (NTAC) to obtain a key or password. NTAC is a Home Office unit in the Crime Reduction and Community Safety group. Within NTAC the Forensic Computing Team (Stored Data) is responsible for providing technical support to UK law enforcement and intelligence agencies in order to assist them to gain access to protected data. Another method of understanding what might be included in encrypted files, in the absence of being able to view the plaintext, is to interpret the encrypted information against known data, as in the US case of *United States of America v Hersh a.k.a. Mario*.[1] In his summary of the facts, Marcus CJ pointed out that a search of Hersh's residence uncovered evidence of computer images of juvenile males engaged in sexual activities. A number of files were encrypted, and the judge described how the investigators dealt with the images as follows:

> Several computer files containing child pornography were found in Hersh's residence: (1) three recovered computer files with viewable images found on the C-drive of Hersh's computer, and (2) encrypted files found on a high-capacity Zip disk. The images on the Zip disk had been encrypted by software known as F-Secure, which was found on Hersh's computer. When agents could not break the encryption code, they obtained a partial source code from the manufacturer that allowed them to interpret information on the file print outs. The Zip disk contained 1,090 computer files, each identified in the directory by a unique file name … that was consistent with names of child pornography files. The list of encrypted files was compared with a government database of child pornography. Agents compared the 1,090 files on Hersh's Zip disk with the database and matched 120 file names. Twenty-two of those had the same number of pre-encryption computer bytes as the pre-encrypted version of the files on Hersh's Zip disk.[2]

1 297 F.3d 1233 (11th Cir. 2002); see also J Alex Halderman, and others, 'Lest we remember: Cold boot attacks on encryption keys' in *Proceedings of the 17th Usenix Security Symposium* (2008) 45, <https://citp.princeton.edu/research/memory/>.
2 297 F.3d 1233, 1238 (11th Cir. 2002) fn 4.

8.8 In this instance, although the files could not be decrypted, there was a sufficient link between the names of the files and evidence of child pornography known to the police.[1] For that reason, the judge drew the inference that the encrypted files contained abusive images of children.

1 For some problems that encryption might cause with the authentication of digital evidence, see Eric Thompson, 'MD5 collisions and the impact on computer forensics' (2005) 2 Digital Investigation 36. For an example where an encoded message was sent by an accused whilst in a county jail awaiting trial and subsequently used to help prove guilt, see Dorn Vernessa Samuel, 'Code breaking in law enforcement: A 400-year history' (2006) 8 Forensic Science Communications.

The UK statutory regime

Notice to require disclosure

8.9 The statutory regime for the investigation of encrypted data is set out in Part III of RIPA 2000. It provides for the investigation of 'protected electronic information', which is defined as 'any electronic data, which, without a key to the data cannot, or cannot readily, be accessed, or be put into an intelligible form.'[1] In 2007, a further Code of Practice, the 'Investigation of Protected Electronic Information Code of Practice',[2] was published pursuant to RIPA to supplement the rules for dealing with encrypted materials. The Code is important because it provides guidance to be followed by any person (other than a judicial authority or a person holding judicial office) when exercising powers under Part III of RIPA 2000 to require the disclosure of protected electronic information in an intelligible form,[3] the means to obtain access to protected information[4] and the means by which protected electronic information may be viewed or put into an intelligible form.[5] Under this regime, authorized persons can, with permission, serve notices on individuals or bodies, to require the disclosure of protected information.

1 Regulation of Investigatory Powers Act 2000, s 53; *Investigation of Protected Electronic Information: Code of Practice* (2007), para 3.12.
2 (The Stationery Office 2007). The Code was issued pursuant to s 71 of the RIPA 2000. Part III of RIPA 2000 covering Protected Electronic Information (Encryption) came into force on 1 October 2007 under the Regulation of Investigatory Powers (Investigation of Protected Electronic Information: Code of Practice) Order 2007 (SI 2007/2200).
3 RIPA 2000, s 49.
4 RIPA 2000, s 50(3)(c).
5 RIPA 2000, s 50(3)(c).

8.10 The relevant power to require disclosure is provided in s 49, RIPA 2000. If 'protected electronic information' has come into the possession of a person by means of the exercise of a statutory authority, or where by any other lawful means not involving the exercise of statutory powers, it has come into the possession of the intelligence services, the police or the customs and excise, such a person may require the disclosure of the key by serving a s 49 notice. Depending on the power under which the protected information was or is likely to be obtained, the persons who have the appropriate permission to serve a s 49 notice are set out in the provisions of Schedule 2 of RIPA.[1]

1 RIPA 2000, s 49(11) and sch 2.

8.11 Other additional criteria must be met before a s 49 disclosure notice may be served. The disclosure must be necessary in the interests of national security, for the purpose of preventing or detecting crime, or be in the interests of the economic well-being of the United Kingdom.[1] The imposition of the notice must be proportionate to

what is sought to be achieved by imposing it,[2] and it must not be reasonably practicable to obtain possession of the protected information in an intelligible form without issuing a notice.[3] Most importantly, the notice to be based on reasonable grounds for believing that a key to the protected information is in the possession of a person.[4]

RIPA 2000, s 49(2)(b)(i) and 49(3). See Yaman Akdeniz, Nick Taylor and Clive Walker, 'Regulation of Investigatory Powers Act 2000: Part 1: BigBrother.gov.uk: State surveillance in the age of information and rights' [2001] Crim LR 73, 85–6 for comments relating to the provisions of s 49(3).
2 RIPA 2000, s 49(2)(c).
3 RIPA 2000, s 49(2)(d).
4 RIPA 2000, s 49(2)(a).

Possession of a key

8.12 The person having possession of information or a key to protected information, is defined in s 56(2), RIPA 2000, which states:

> References in this Part to a person's having information (including a key to protected information) in his possession include references-
>
> (a) to its being in the possession of a person who is under his control so far as that information is concerned;
>
> (b) to his having an immediate right of access to it, or an immediate right to have it transmitted or otherwise supplied to him; and
>
> (c) to its being, or being contained in, anything which he or a person under his control is entitled, in exercise of any statutory power and without otherwise taking possession of it, to detain, inspect or search.

8.13 This definition postulates three scenarios for possessing a key:

> (i) a person may possess a key if it is under his control, or
>
> (ii) if he has an immediate right of access to it, or an immediate right to have it transmitted or supplied to him, or
>
> (iii) if he (or a person under his control) is entitled, in exercise of any statutory power and without taking possession of it, to detain, inspect or search the thing which contains the key.

8.14 In the second and third scenarios, a person may be deemed to have a key, although he does not have the key himself. This is a fairly important provision, because the managerial officers of an organization, whatever the legal form the organization takes, are the ones responsible for the proper management of the private key, rather than the operational staff members.[1] Thus any s 49 notice should be served on an officer or senior manager of the organization.

1 See Ross Anderson, *Security Engineering: A Guide to Building Dependable Distributed Systems* (2nd edn, John Wiley & Sons 2008) para 3.7.4 for a discussion on the principles involved in this process.

Notice requiring disclosure

8.15 The format a disclosure notice must take is set out in s 49(4), RIPA 2000. It must, among other things, describe the protected information to which the notice relates;[1] specify the grounds upon which the disclosure is believed to be necessary;[2] specify the time by which the notice is to be complied with,[3] which must allow a reasonable period for compliance, depending on the circumstances of the case,[4] and specify the

disclosure required and the form and manner in which it is to be made.[5] The required disclosure could be the protected information in an intelligible form,[6] or the disclosure of the key to the protected information.[7] Where there is a cost to complying with a s 49 notice, s 52 provides for the Secretary of State to make an appropriate contribution towards such costs.

1 RIPA 2000, s 49(4)(b).
2 RIPA 2000, s 49(4)(c).
3 RIPA 2000, s 49(4)(f).
4 RIPA 2000, s 49(4) proviso.
5 RIPA 2000, s 49(4)(g).
6 RIPA 2000, s 50(1).
7 RIPA 2000, s 50(2), (4)–(7).

Disclosure of protected information and keys

8.16 Where a person is served with a s 49 notice requiring the disclosure of protected information in an intelligible form, he may use any key in his possession to obtain access to the information or to put it into intelligible form.[1] However, the person to whom the notice is addressed may instead disclose the relevant key to the person serving the notice.[2] There may be times when the person to whom the notice is directed does not have the key, or cannot gain access to the key. In such instances, he must give up what keys he actually has, although he does not have to disclose every key he has in his possession.[3]

1 RIPA 2000, s 50(1).
2 RIPA 2000, s 50(2).
3 RIPA 2000, s 50(3) and the effects of s 50(4), (5) and (6). See also s 50(7) and (8).

8.17 It follows that where a notice is to be served on a body corporate or a firm and it is obvious that more than one person may be in possession of the key, then the notice will be directed to a senior officer, partner or senior employee.[1] However, where it is considered that the circumstances are such that the purpose of the notice would be defeated if it were to be served on the most appropriate person (for instance, he may be the subject of an investigation), then the notice may be served on another individual.[2]

1 RIPA 2000, s 49(5) and (6).
2 RIPA 2000, s 49(7).

8.18 An exception is created as regards the disclosure of keys that are used for generating electronic signatures. Section 56(1) RIPA 2000 defines an 'electronic signature' as:

> anything in electronic form which
>
> (a) is incorporated into or logically associated with, any electronic communication or other data;
>
> (b) is generated by the signatory or other source of the communication or data; and
>
> (c) is used for the purpose of facilitating, by means of a link between the signatory or other source and the communication or data, the establishment of the authenticity of the communication or data, the establishment of its integrity, or both;

8.19 Where a key is used only for this purpose, it does not have to be disclosed in response to a notice, provided it has in fact not been used for any other purpose.[1] It might be useful to recall that a key pair has more than the single function of producing an electronic signature. The same key pair can be used to encrypt a message, depending on the algorithm used.

1 RIPA 2000, s 49(9).

8.20 However, this exemption may be narrower than it seems. In a commercial context, where more than one person may properly have access to a key, the person served with the notice may not be able to be sure that a key, despite being intended for signature purposes, has never been used to decrypt a message encrypted with the corresponding public key (there is no disclosure obligation if the key 'has not in fact been used for any … purpose [other than that of generating electronic signatures]').[1] Although it is arguably for the prosecution to prove that a key has been used for such a purpose (and is therefore subject to seizure), the mere assertion of this fact by the person demanding access to the key would place the recipient of the notice in a difficult position to prove a negative in resisting the demand.

1 RIPA 2000, s 49(9)(b).

When the key is required

8.21 There may be occasions when the key is required, rather than the unencrypted plain text.[1] In such circumstances, the provisions of s 51, RIPA 2000, will apply. The criteria are set out in s 51(4), in that a direction to produce the key cannot be given unless it is believed that the circumstances are such that the purpose would be defeated if the notice did not provide for the key to be delivered up, and that asking for the key is proportionate. When deciding whether the demand for the key is proportionate, s 51(5) sets out the matters to be taken into account, including the nature and extent of the protected information and any adverse effects that asking for the key might have on the business carried on by the person to whom the notice is directed.

1 RIPA 2000, s 50(3)(c).

Circumventing a notice

8.22 It has been held, albeit at first instance, in an unreported decision, that the procedure set out by the RIP Act 2000 is the only way to compel the disclosure of the key to an encrypted device. In this case, Laurie Love was accused of hacking into computers in the USA. The National Crime Agency (NCA) executed a search warrant and seized his computers and devices. Some of the devices contained encrypted files and permission was sought, and received, to serve a s 49 notice. Love refused to comply with the order. He was not prosecuted for this.

8.23 The USA began proceedings to extradite him, and Love also began proceedings under s 1 of the Police Property Act 1897 to have his computers returned to him. The District Judge, hearing the case, initially ordered Love to file witness statements to explain what data were stored on the various encrypted devices. Love refused to accede to that order. The NCA then applied to the court for an order that the encryption key or password to the encrypted items should be disclosed before the property was returned to Love. District Judge Tempia, held that he could not do so.[1] The power

to make directions in such cases is to be found in the Criminal Procedure Rules and are best described as case-management powers. The District Judge held that case management powers could not be used to circumvent a statutory process that had been created for a specific purpose.

1 *Laurie Love v National Crime Agency*, unreported, 2 March 2016, City of Westminster Magistrates' Court (case 011503187270).

8.24 That must be correct. Parliament put forward a scheme to deal with encrypted data, including a series of safeguards. Indeed the NCA had served a s 49 notice which Love refused to comply with. The correct response to such refusal is to prosecute Love under s 53 (where the police and prosecutors believe there was no reasonable excuse for not complying). However, that an applicant refuses to disclose the keys must inevitably be something that the courts can take into account when deciding to return property. If a person is suspected of hiding illegal activity through the use of encryption and refuses to disclose the contents, it is unlikely that he will succeed in an application to have his property returned, because to do so would potentially mean returning inculpatory evidence to him (and which may lead to the ultimate destruction of the only evidence against him).

Failure to comply with a notice

8.25 A person to whom a s 49 notice has been given is guilty of an offence in s 53, RIPA 2000, if he knowingly fails to make the disclosure required by the notice.[1] The provisions of this section are important. The presumption of possession of a key to protected information is set out in s 53(2) and (3), RIPA 2000:

> (2) In proceedings against any person for an offence under this section, if it is shown that that person was in possession of a key to any protected information at any time before the time of the giving of the section 49 notice, that person shall be taken for the purposes of those proceedings to have continued to be in possession of that key at all subsequent times, unless it is shown that the key was not in his possession after the giving of the notice and before the time by which he was required to disclose it.
>
> (3) For the purposes of this section a person shall be taken to have shown that he was not in possession of a key to protected information at a particular time if—
>
> (a) sufficient evidence of that fact is adduced to raise an issue with respect to it; and
>
> (b) the contrary is not proved beyond a reasonable doubt.

1 RIPA 2000, s 53(1) and (5).

8.26 The provisions of this section have the following effect: the prosecution has the persuasive burden to prove that the accused was in possession of a key to protected information at any time before the giving of the s 49 notice.[1] This element of the definition of the offence requires the prosecution to prove possession, not merely that a key-generating third party sent the key. Also, the second element relates to time. The key has to be proved to be in the possession of the accused at any time up to the giving of the s 49 notice. This part of the offence means that it is possible for a key to have expired and to have been deleted, if such was the policy in a commercial organization, for instance. It has been argued that the presumption that the key remains with the

person, as a continuing state of affairs, may be unfair.[2] In any event, the key management policy becomes an important document, as does the physical implementation of its provisions, in any prosecution for failure to comply with a disclosure notice.

1 It is possible for encrypted data to be encoded in such a way that it can be decoded in two separate ways, one to reveal the secret message and the other to reveal an innocuous message: Derrick Grover, 'Dual encryption and plausible deniability' (2004) 20 Computer L & Secur Rep 37; Derrick Grover, 'Data – plausible deniability' (2005) 21 Computer L & Secur Rep 405; many freely available encryption software programs include plausible deniability options. The implications of this are significant because, if used properly, it would be very difficult to prove forensically that there were, in fact, two ways of decoding the encrypted space. A forensic examination would not show two volumes, only one. Where a user does not clear document histories and such like, it may be possible to show the existence of a second volume, but otherwise it would be very difficult.

2 Yaman Akdeniz, Nick Taylor and Clive Walker, 'Regulation of Investigatory Powers Act 2000: Part 1: BigBrother.gov.uk: State surveillance in the age of information and rights' [2001] Crim LR 73, 87.

8.27 The second part to the offence places an express burden on the accused to offer a reason for why a key is not in his possession: 'unless it is shown that the key was not in his possession after the giving of the notice and before the time by which he was required to disclose it.' The provisions of s 53(3) offer further guidance relating to the defence the accused is permitted to raise:

> (3) For the purposes of this section a person shall be taken to have shown that he was not in possession of a key to protected information at a particular time if-
>
> (a) sufficient evidence of that fact is adduced to raise an issue with respect to it; and
>
> (b) the contrary is not proved beyond a reasonable doubt.

8.28 Section 53(3) provides that the defendant has the burden of proving that there is some factual evidence to suggest he may not be in possession of the key. It will then be for the prosecution to disprove it beyond all reasonable doubt. This is a common way of establishing reverse burdens. The prosecution cannot be expected to rebut all potential issues. However, this does not require, for example, the defendant to prove a matter on the balance of probabilities, but simply to adduce some evidence to show that it is a live issue.[1] That said, the quality of the evidence given will be important in deciding whether it is, indeed, a live issue.

1 The leading authority on reverse burdens remains *R v Lambert* [2001] UKHL 37.

Sentencing

8.29 To date, the courts have only considered the sentencing for offences under s 53 (failing to comply with a notice), but there are some interesting points arising from this which are worthy of discussion, because it can be relevant to both an investigator and prosecutor.

8.30 Section 53(5) provides that where the case is a national security or child indecency case, the maximum sentence is five years' imprisonment; otherwise it is two years. A national security case is one in which 'the grounds specified in the notice to which the offence relates as the grounds for imposing a disclosure requirement were or included a belief that the imposition of the requirement was necessary in the interests of national security'.[1] A child indecency case is one in which 'the grounds specified

in the notice to which the offence relates as the grounds for imposing a disclosure requirement were or included a belief that the imposition of the requirement was necessary for the purposes of detecting an offence under any of the provisions listed under subsection (7)'.[2] The offences listed in s 53(7) are those that relate to the possession, making, taking or distribution of indecent photographs of children in England & Wales, Scotland and Northern Ireland.

1 RIPA 2000, s 53(5B).
2 RIPA 2000, s 53(6).

8.31 What is interesting about the sentence is that it relies on the belief of the investigator. The sentence is based on the grounds that are mentioned in the disclosure notice. It will be remembered that the standard of proof for such applications – and therefore the grounds for application – is reasonable belief, a standard significantly below the usual standard for criminal offences. Of course, a judge who hears the application and the judge who tries the s 53 offence (who should be different) could consider the reasonableness of the belief in deciding whether to grant permission or, in relation to s 53, whether the prosecution can take place.[1] Notwithstanding this, it is perhaps surprising that the prosecution does not need to prove that it is a national security or child indecency case. While it would be very difficult to prove this to the standard of beyond all reasonable doubt (as the protected information would be required for this), it may be possible to show this on the balance of probabilities (through circumstantial evidence, such as downloading history, Internet search terms, IP logs etc.). In any case, it should not suffice for the prosecution to simply demonstrate a 'reasonable belief' for such a fundamental aspect of the prosecution's case for a serious offence. In *R v Cutler*[2] the Court of Appeal explained this offence thus:

> [A s 53 offence is] a very serious offence because it interferes with the administration of justice and it prevents the prosecuting authorities and the police finding out what offences someone has committed.[3]

1 Where it was found that the police had artificially considered a case a national security or child indecency case it is inevitable that this would be considered an abuse of process and the prosecution would be stayed.
2 [2011] EWCA Crim 2781.
3 [2011] EWCA Crim 2781, [35].

8.32 This is an important point. As has been noted, encryption puts evidence beyond the reach of law enforcement and prosecutors. It means the full extent of the criminality cannot be ascertained, and the courts must consider this seriously. It is, if s 53 is proven, a deliberate attempt to try and conceal evidence from the competent authorities, and this must merit harsh sanctions.

8.33 The seriousness of the offence is perhaps reflected in the comments of the Court of Appeal in *R v Padellec.*[1] The appellant entered a plea of guilty to an offence under s 53. He came to the attention of the police as a possible acquaintance of a person known to be involved in the trafficking of children. His computer (which included an encrypted folder) was recovered, and while no indecent images of children were found, search terms relating to indecent photographs were found. The appellant alleged that he purchased the encrypted device in Belgium and had no knowledge of the key. Following negotiations, a basis for the plea was tendered and accepted by the Crown. This was as follows:

> 1. The defendant accepts that he did not provide passwords as requested.
>
> 2. He did not do so because he knew he had used wiping software to remove evidence of a small number of images, which he accepts were indecent.
>
> 3. The defendant had accessed these images during the currency of Internet browsing. The defendant will assert that the content of these images did not depict images of very young children. He cannot state the ages. The images did not contain scenes of sexual or any other type of violence to children.[2]

1 [2012] EWCA Crim 1956.
2 [2012] EWCA Crim 1956, [6].

8.34 The importance of the third basis of plea is that it, in essence, the defendant did not obtain access to the images of the very worst forms of indecent photographs of children.[1] The judge accepted the plea, but suggested that he did it with reluctance. The Court of Appeal was scathing about the basis of plea. In giving judgment, Collins J said:

> It seems to us that in a case such as this, it is entirely wrong for a basis of plea to be accepted, either by the prosecution or ultimately by the judge. What it does is to enable the defendant in question to identify, to his advantage, what was or was not on the computer and to get a lesser sentence than otherwise might be appropriate. That is to enable him to dictate, wrongly, what the situation is. The whole point of requiring access is so that it can be seen what was, in fact, there. We express hope that in a situation that arose in this case, there will never again be a basis of plea accepted which is based on keeping the contents secret and the defendant saying, to his advantage, what was or was not contained.[2]

1 At the time of this decision, the sentencing of indecent photographs was subject to the definitive sentencing guideline of 2007. This created five categories of seriousness. The basis of plea would ensure that it did not fall within the highest category or contain any aggravating factors. The guideline was replaced in 2013, but the changes are irrelevant to this decision.
2 [2012] EWCA Crim 1956, [11].

8.35 If the defendant had not viewed, or stored, images that constituted the most serious examples of indecent photographs of children, then he could have proved this by allowing access to the device. Instead, the prosecution (and the judge) decided that the defendant could admit that he had looked at illegal content but could also keep the details of this illegality secret. The Court of Appeal, quite rightly, considered this an affront to justice. The judge (rightly) believed that in the absence of an explanation, an assumption should be made that a person is sentenced on the worst-case basis. To avoid this, the defendant need only prove what he had admitted to. However, that requires proof, and it is quite wrong to reward an offender in the form of discounting a sentence by trusting what he said when his veracity could be shown by decrypting the device.

Obligations of secrecy and tipping off

8.36 There is a power to attach a secrecy provision to any disclosure requirement.[1] This will require the person to whom the notice is given, and every other person who becomes aware of its contents, to keep the giving of the notice, its contents and the things undertaken in responding to it, a secret.[2] It is a criminal offence where a person fails to comply with a disclosure requirement or a secrecy requirement. The penalty depends on what is believed to be hidden behind the key. Where the material

is suspected to relate to child pornography or a national security case, the maximum sentence is five years' imprisonment,[3] whereas for all other cases it is two years' imprisonment.[4] The criteria for the imposition of a secrecy requirement are set out in s 54(3), and the punishment is provided in s 54(4). A range of defences are set out in ss 54(5)–(10).

1 RIPA 2000, s 54.
2 RIPA 2000, s 54(1).
3 RIPA 2000, s 53(5A)(a).
4 RIPA 2000, s 53(5A)(b).

8.37 It should be noted, however, that the effectiveness of the 'tipping off' offence is debatable. It might be possible for a person to sign his email correspondence with a disclaimer, such as 'I will always explain why I revoke a key, unless the UK government prevents me using the RIP Act 2000'. Using this qualification, let it be assumed that a correspondent revokes a key. If the correspondent is asked for the reason and he replies that he cannot give one, it is doubtful if he can be convicted of the offence of tipping off, though this is exactly what he has done. There is no suggestion that a disclosed key cannot lawfully be revoked.

Refusal to reveal the key

8.38 Most suspects will not offer up the relevant password or key voluntarily, and may mount various challenges to the s 49 disclosure notice. In *R v S (F) and A (S)*,[1] the defendants challenged the validity of s 49 notices served on them. The facts were as follows. In 2007, H was made the subject of a control order under the Prevention of Terrorism Act 2005. He was required to live and remain in Leicestershire, and not to leave his home address without the consent of the Secretary of State for the Home Department. The appellants are alleged to have conspired together, and with H and others, to breach the order, by helping H abscond from his address in Leicester and conveying him to a secret address in Sheffield, which S did on 9 September 2007. Shortly after their arrival, the police entered the premises. H and S were found in separate rooms. S was alone in the same room as a computer. The password to an encrypted file appeared to have been partially entered. He was arrested, subsequently interviewed, and made no comment. A search of his home address in London revealed computer material. A number of documents had been deleted from the computer hard drives, but when retrieved, they provided the basis for charges against S under s 58 of the Terrorism Act 2000. However, in the absence of the passwords for the encrypted files present on the computer hard drives, and the full password for an encrypted file on the laptop upon which the encryption key appeared to have been already partially entered in Sheffield, the encrypted files could not be opened. A was also arrested on 9 September 2007. His address was also searched, and the police seized computer material. One of the disks seized included an area on the disk that was encrypted.

1 [2008] EWCA Crim 2177, [2009] 1 All ER 716, [2009] 1 WLR 1489; see also *R v Cutler* [2011] EWCA Crim 2781, 2011 WL 5902910.

8.39 S and A were each served with a notice under s 53 of RIPA 2000. Neither complied with the notices, and argued, in essence, that the notices compelling them to

disclose the passwords or the keys to the encrypted computer files were incompatible with the privilege against self-incrimination.

8.40 It is a criminal offence knowingly to refuse or fail to make the disclosure required by a notice issued under s 49.[1] The data that was encrypted might have contained incriminating information, but it was not certain that it would contain incriminating information. During the course of a preparatory hearing on 26 June 2008, HHJ Stephens QC decided that the privilege against self-incrimination was not available because the encrypted material existed, which meant its existence did not depend on the appellants, and the notice was legitimate and proportionate. After briefly discussing the privilege against self-incrimination, and the limits that apply to the privilege, the President raised the question as to whether the principle itself was engaged in each individual case.[2] The arguments concentrated on whether the passwords to the keys were properly a piece of information with an existence separate from the 'will' of each appellant. The President's analysis is set out as follows:

> On analysis, the key which provides access to protected data, like the data itself, exists separately from each appellant's "will". Even if it is true that each created his own key, once created, the key to the data, remains independent of the appellant's 'will' even when it is retained only in his memory, at any rate until it is changed. If investigating officers were able to identify the key from a different source (say, for example, from the records of the shop where the equipment was purchased) no one would argue that the key was not distinct from the equipment which was to be accessed, and indeed the individual who owned the equipment and knew the key to it. Again, if the arresting officers had arrived at the premises in Sheffield immediately after S had completed the process of accessing his own equipment enabling them to identify the key, the key itself would have been a piece of information existing, at this point, independently of S himself and would have been immediately available to the police for their use in the investigation. In this sense the key to the computer equipment is no different to the key to a locked drawer. The contents of the drawer exist independently of the suspect: so does the key to it. The contents may or may not be incriminating: the key is neutral. In the present cases the prosecution is in possession of the drawer: it cannot however gain access to the contents. The lock cannot be broken or picked, and the drawer itself cannot be damaged without destroying the contents.[3]

1 RIPA 2000, s 53 is discussed below.

2 The privilege against self-incrimination was considered by the European Court of Human Rights in *Saunders v United Kingdom* [1997] BCC 872, [1998] 1 BCLC 362, (1997) 23 EHRR 313, where Mr Saunders was compelled by a statutory power to give evidence to DTI inspectors, and the evidence was later used in his criminal trial. The court stated at para [69], that: 'As commonly understood in the legal systems of the Contracting Parties to the Convention and elsewhere, [the privilege against self-incrimination] does not extend to the use in criminal proceedings of material which may be obtained from the accused through the use of compulsory powers but which has an existence independent of the will of the suspect …'

3 [2008] EWCA Crim 2177, [20].

8.41 In this case, only the appellants had the passwords to decrypt the documents. The President noted that if they gave up the passwords as required, 'The actual answers, that is to say the product of the appellants' minds could not, of themselves, be incriminating. The keys themselves simply open the locked drawer, revealing its contents'.[1] The President continued his analysis:

> If however, as for present purposes we are assuming, they contain incriminating material, the fact of the appellants' *knowledge* of the keys may itself become an incriminating fact. For example, to know the key to a computer in your possession which contains indecent images of children may itself tend to support the prosecution case that you were knowingly in possession of such material.[2]

1 [2008] EWCA Crim 2177, [21].
2 [2008] EWCA Crim 2177, [21].

8.42 A distinction was made as to the circumstances where knowledge of the password will be relevant to the privilege against self-incrimination. The privilege would only apply if the data, which exists independently of the will of the appellants (the privilege against self-incrimination does not apply to the data), contains incriminating material. If the data did not contain incriminating material, then the knowledge of how to obtain access to it would also not be incriminating. The President continued:

> … the question which arises, if the privilege is engaged at all, is whether the interference with it is proportionate and permissible. A number of issues are clear and stark. The material which really matters is lawfully in the hands of the police. Without the key it is unreadable. That is all. The process of making it readable should not alter it other than putting it into an unencrypted and intelligible form that it was in prior to encryption; the material in the possession of the police will simply be revealed for what it is. To enable the otherwise unreadable to be read is a legitimate objective which deals with a recognised problem of encryption. The key or password is, as we have explained, a fact. It does not constitute an admission of guilt. Only knowledge of it may be incriminating. The purpose of the statute is to regulate the use of encrypted material, and to impose limitations on the circumstances in which it may be used. The requirement for information is based on the interests of national security and the prevention and detection of crime, and is expressly subject to a proportionality test and judicial oversight. In the end the requirement to disclose extends no further than the provision of the key or password or access to the information. No further questions arise. The notice is in very simple form. Procedural safeguards and limitations on the circumstances in which this notice may be served are addressed in a comprehensive structure, and in relation to any subsequent trial, the powers under section 78 of the 1984 Act to exclude evidence in relation, first, to the underlying material, second, the key or means of access to it, and third, an individual defendant's knowledge of the key or means of access, remain. Neither the process, nor any subsequent trial can realistically be stigmatised as unfair.[1]

1 [2008] EWCA Crim 2177, [25].

8.43 Roberts, in the Criminal Law Review, observed that an encryption key, unless documented, is an 'intangible "psychological fact", that is to say, it is information which exists only in the suspect's memory and that of any other person who might "know" it'.[1] Encryption keys are usually far too long for any person to commit to memory.[2] A person can only encrypt data with an encryption key, but the encryption key itself is usually protected by a further form of encryption, possibly by way of an encryption application, which in turn is protected by a password. It is the password that is the 'intangible' item of knowledge – that is, it will be intangible if it is not recorded – not the encryption key. In this respect, Roberts is correct to note that the password that controls the encryption key cannot be distinguished from the knowledge of the accused.

1 [2009] Crim LR 191, 192; in 1993, Professor Tapper observed that the increased use of computers will lead to the position that we recess 'to the earlier period where information reposed only in the brains of those who were party to it, and had no material form': Colin Tapper, 'Evanescent evidence' (1993) 1 Intl J L & Info Tech 35, 40.
2 For an example of what a private key looks like, see Stephen Mason, *Electronic Signatures in Law* (4th edn, University of London 2016) 153.

8.44 If the encrypted data were revealed, and they contained incriminating material, then it would be for the trial judge to exclude the evidence under s 78 of the Police and Criminal Evidence Act 1984 if it is considered appropriate. The Court of Appeal upheld the decision of HHJ Stephens QC because the purpose of the statute is to regulate the use of encrypted material and to impose limitations on the circumstances in which it may be used, subject to a proportionality test and judicial oversight, and neither the process, nor any subsequent trial could be considered to be unfair.[1]

1 In May 2009, Oliver Drage, 19, of Liverpool, was arrested by police officers investigating child sexual exploitation. His computer was seized. It was protected by a 50-character password. He was convicted of failing to disclose an encryption key in September 2010. He was sentenced to 16 weeks' imprisonment at Preston Crown Court on 4 October 2010: 'Man jailed over computer password refusal', *BBC News* (5 October 2010) <www.bbc.co.uk/news/uk-england-11479831>; for an example of a sentence, see *R v Padellec* [2012] EWCA Crim 1956.

8.45 Another case in which the issues were aired was the case of *Greater Manchester Police v Andrews*,[1] an appeal from the refusal of HHJ Steiger QC to issue an order to serve a s 49 notice on Anthony Andrews. Andrews had been convicted of sexual offences committed on two young girls and one young boy, and was subsequently arrested on suspicion of breaching a Sexual Offences Prevention Order. The police seized his laptop computer and two memory sticks. An examination of the computer revealed indecent images of children. The memory sticks were encrypted. During interview, Andrews declined to answer any questions, including questions in relation to the passwords and software applications that had been used to prevent access to the files. As a result, the police applied to serve a notice on Andrews under s 49 of RIPA 2000, requiring disclosure of the encryption keys. HHJ Steiger QC refused that application, on the basis that for Andrews to reveal the key would risk his privilege against self-incrimination, as there was no other independent evidence to show that he did know what the key was,[2] and on that basis, distinguished the case of *R v S (F) and A (S)*.[3]

1 [2011] EWHC 1966 (Admin), [2012] ACD 18.
2 [2011] EWHC 1966 (Admin), [18]–[19].
3 [2008] EWCA Crim 2177, [2009] 1 All ER 716, [2009] 1 WLR 1489.

8.46 On the morning of the appeal hearing, the court was informed that since these events, Andrews had been rearrested on suspicion of further similar offences involving more abusive images of children, quite separate from the images found on the laptop seized by the police. He entered a plea of guilty to a number of offences, and was sentenced to a term of imprisonment for public protection with a minimum custodial term of 27 months. The court continued with the appeal on the basis that the matter was of real public interest in the protection of prevention of crime because of the possibility of the encrypted material potentially disclosing either the victim of or perpetrators of this type of offence.

8.47 In giving judgment, McCombe J considered that HHJ Steiger took 'an extremely limited view of the evidence before the court as to what the respondent's knowledge of the key was. To my mind it was a perfectly legitimate inference to draw from the circumstances of recovery of the pen drives that the respondent might know encryption keys relating to the information stored on them.'[1] The members of the court agreed that the privilege against self-incrimination was engaged in this case, but only to a very limited extent, and it was proportionate and in the public interest within the meaning of s 49(2)(c) for Andrews to be required to give up the key. Sir Anthony May P agreed, and observed:

> Privilege against self-incrimination is not absolute and it is plain that this statute does not intend that it should be. Section 49(2)(c) requires that the imposition of a disclosure requirement has to be proportionate to what is sought to be achieved. Since the nature of the disclosure is very likely to be concerned with criminal activity it is implicit from this that there may be circumstances in which it is proportionate to require disclosure even though the privilege against self-incrimination may arise for consideration to be a very limited extent.[2]

1 [2011] EWHC 1966 (Admin), [21].
2 [2011] EWHC 1966 (Admin), [27].

The approach in the United States of America

8.48 These cases may be usefully contrasted with the approach taken in the United States of America. In particular, the case of *In re Grand Jury Subpoena to Sebastien Boucher*[1] before the U.S. District Court for the District of Vermont was cited in *R v S (F) and A (S)*[2] to illustrate the point that knowledge of the password might be relevant to the privilege against self-incrimination. This is one of the first cases in which this problem arose. The facts were that on 17 December 2006, Boucher and his father entered the United States from Canada. A Customs and Border Protection Officer found a laptop computer in the vehicle they were travelling in. He opened the computer and switched it on without entering a password. He searched the various files in the computer, and discovered approximately 40,000 images, some of which appeared to be pornographic, based on the names of the files. Boucher was asked if any of the files contained abusive images of children, to which he responded that he was not certain. The officer continued to search the files, and noticed some files with names that suggested child pornography. He then requested the help of another officer, who determined that a number of files contained abusive images of children. Boucher was then read his *Miranda* rights. He told the second officer that he downloaded pornographic files, and indicated that he did not intentionally download child pornography and deleted any such images when he came across them. Boucher was given access to the laptop and navigated to Z drive, which he obtained access by inserting a password. The second officer did not see Boucher do this. Boucher was subsequently arrested and his laptop was seized. After obtaining a search warrant, the government discovered that the Z drive was encrypted. The investigating authorities could not open the Z drive. A grand jury subpoena was issued for Boucher, directing him to:

> … provide all documents, whether in electronic or paper form, reflecting any passwords used or associated with the Alienware Notebook Computer … [model and serial numbers] … seized from … Boucher at [place and date][3]

1 2007 WL 4246473 (D.Vt.).

2 [2008] EWCA Crim 2177, [2008] WLR (D) 313, [2010] Crim LR 191.
3 2007 WL 4246473 (D.Vt.), [2].

8.49 Boucher moved to quash the subpoena because, he alleged, it violated his right not to incriminate himself under the provisions of the Fifth Amendment. Whether the privilege against self-incrimination applied in this instance depended on whether the subpoena sought testimonial communication.[1] Both parties agreed that the contents of the laptop computer were not covered by the Fifth Amendment, because they were voluntarily prepared and not testimonial in nature. Niedermeier MJ, commented that:

> Entering a password into the computer implicitly communicates facts. By entering the password Boucher would be disclosing the fact that he knows the password and has control over the files on drive Z. The procedure is equivalent to asking Boucher, "Do you know the password to the laptop?" If Boucher does know the password, he would be faced with the forbidden trilemma; incriminate himself, lie under oath, or find himself in contempt of court.[2]

1 See *State of Florida v Stahl*, 206 So.3d 124 (2016), 2016 WL 7118574, 41 Fla. L. Weekly D2706 (requiring the defendant to produce a passcode did not compel the defendant to communicate information that had testimonial significance).
2 2007 WL 4246473 (D.Vt.), [3].

8.50 The judge concluded that the provisions of the Fifth Amendment prevented the government from compelling Boucher from providing the password on the basis that it would compel him to display the contents of his mind to incriminate himself:

> While the government may know of the existence and location of the files it has previously viewed, it does not know of the existence of other files on drive Z that may contain incriminating material. By compelling entry of the password the government would be compelling production of all the files on drive Z, both known and unknown.
>
> …
>
> The password is not a physical thing. If Boucher knows the password, it only exists in his mind. This information is unlike a document, to which the foregone conclusions doctrine usually applies, and unlike any physical evidence having testimonial aspects. Compelling Boucher to produce the password compels him to display the contents of his mind to discriminate himself.[1]

1 2007 WL 4246473 (D.Vt.), [6].

8.51 The government appealed this decision,[1] arguing that the government was already aware of the existence and location of the information during the border examination (when the officer viewed the contents of some of the Z drive files, and ascertained that they could consist of images or videos of child pornography). Chief District Court Judge William K. Sessions, III agreed, overruling the initial ruling and sustaining the government's appeal. He stated that requiring Boucher to 'provid[e] access to the unencrypted Z drive "adds little or nothing to the sum total of the Government's information" about the existence and location of files that may contain incriminating information,' and therefore this did not constitute 'compelled testimonial communication' and did not breach Boucher's Fifth Amendment right against self-incrimination.[2]

1 *In re Grand Jury Subpoena to Sebastien Boucher*, 2009 WL 424718 (D.Vt.).
2 *In re Grand Jury Subpoena to Sebastien Boucher*, 2009 WL 424718 (D.Vt.), [2]-[3]. For more

discussion in the US context and reference to other articles, see Aaron M Clemens, 'No computer exception to the constitution: The Fifth Amendment protects against compelled production of an encrypted document or private key' (2004) 8 UCLA Journal of Law and Technology 1; Andrew J Ungberg, 'Protecting privacy through a responsible decryption policy' (2009) 22 Harv J L & Tech 537; John Duong, 'The Intersection of the Fourth and Fifth Amendments in the context of encrypted personal data at the border' (2009) 2 Drexel Law Review 313; David Colarusso, 'Heads in the cloud, A coming storm: The interplay of cloud computing, encryption, and the Fifth Amendment's protection against self-incrimination' (2011) 17 Boston University Journal of Science and Technology Law 69; Adam M Gershowitz, 'Password protected? Can a password save your cell phone from a search incident to arrest?' (2011) 96 Iowa L Rev 1125; Susan W Brenner, 'The Fifth Amendment, cell phones and search incident: A response to password protected?' (2011) 96 Iowa L Rev 78; Michael Wachtel, 'Give Me Your Password Because Congress Can Say So: An Analysis of Fifth Amendment Protection Afforded Individuals Regarding Compelled Production of Encrypted Data and Possible Solutions to the Problem of Getting Data from Someone's Mind' (2013) 14 U Pitt J Tech L & Policy 44; Andrew T Winkler, 'Password protection and self-incrimination: applying the fifth amendment privilege in the technological era' (2013) 39 Rutgers Computer & Tech LJ 194.

8.52 A similar approach was taken by the United States Fourth Circuit in the case of *United States of America v Gavegnano*,[1] in which the appellant was convicted of receipt and possession of abusive images of children stored on a laptop computer owned by the government and issued to him for the purposes of his work. One of the grounds of appeal was based on the Fifth Amendment, in that he gave the password of the laptop computer to the prosecuting authorities after meeting with his lawyer. The members of the Court of Appeals rejected his claim, on the basis that 'Any self-incriminating testimony that he may have provided by revealing the password was already a "foregone conclusion" because the Government independently proved that Gavegnano was the sole user and possessor of the computer.'[2]

1 305 Fed.Appx. 954 (4th Cir. 2009), 2009 WL 106370.
2 305 Fed.Appx. 954, 956 (4th Cir. 2009).

8.53 The *Boucher* case was not cited in the case of *United States of America v Kirschner*,[1] where Borman DJ of the Eastern District of Michigan decided that the subpoena requiring the defendant to give up the password must be quashed on the basis that the government was not seeking documents or objects, but testimony from the defendant by requiring him to divulge – through his mental processes – his password that would be used to incriminate him.[2]

1 2010 WL 1257355 (E.D.Mich.).
2 2010 WL 1257355 (E.D.Mich.), [4].

8.54 In contrast, in the District of Colorado case of *United States v Ramona Camelia Fricosu a/k/a/ Ramona Smith*,[1] Blackburn DJ ordered (having cited *Boucher* and *Kirschner*) that the accused provide the government with an unencrypted copy of the hard drive of a laptop computer found in her bedroom, having established on the preponderance of evidence that the laptop either belonged to the accused, or that she was the sole or primary user, such that she could obtain access to the contents that were encrypted.

1 2012 WL 182121 (D.Colo.).

8.55 In the case of *In re Grand Jury Subpoena Duces Tecum Dated March 25, 2011*,[1] law enforcement agents began an investigation in March 2010 of an individual

suspected of using a YouTube.com account for sharing explicit materials involving underage girls. During the course of their investigation, officers from the Santa Rosa County Sherriff's office in Florida obtained several Internet protocol (IP) addresses from which the individual had obtained access to the Internet. Three of the addresses were subsequently traced to hotels. A review of the register in each hotel revealed a common name registered at the hotel at the relevant time, being that of one Doe. Doe was found at a hotel in California, and the police applied for and obtained a warrant to search his room. Seven items were seized, including two laptops and five external hard drives. Examiners from the Federal Bureau of Investigation analysed the digital media, but could not obtain access to some parts of the hard drives because they were encrypted with a software program called TrueCrypt.

1 670 F.3d 1335 (11th Cir. 2012).

8.56 Doe refused to provide the passwords to enable the government to open and view the encrypted data, and he also refused to decrypt the data. As a result, he was served with a subpoena duces tecum, requiring him to appear before a grand jury and produce the plain text of the encrypted files located on the hard drives of his laptop computers and the five external hard drives. Federal prosecutors offered him immunity for the act of decrypting the computer, but reserved the right to use any evidence it found on the computer against him.[1] When he appeared before the jury, he invoked his right under the Fifth Amendment against self-incrimination not to reveal the plain text. During the hearing, the forensic examiner testified that he could obtain access to some parts of the hard drives, but he could not know for certain whether there might be data on the encrypted part of the hard drive – indeed, he accepted there might not be any data in the encrypted part of the drives. Collier CJ determined that Doe's failure to decrypt the relevant parts of hard drives amounted to contempt of court and committed him to custody. He appealed, and he was released when his appeal was allowed.

1 670 F.3d 1335, 1350 (11th Cir. 2012).

8.57 The Eleventh Circuit Court of Appeals allowed the appeal on two grounds. First, it held that production of the plain text would constitute testimony and was not merely a physical act. The court indicated that what was in issue was whether the act of production may have some testimonial quality sufficient to bring the Fifth Amendment into play. Tjoflat J stated that:

> … the decryption and production of the hard drives would require the use of the contents of Doe's mind and could not be fairly characterized as a physical act that would be nontestimonial in nature. We conclude that the decryption and production would be tantamount to testimony by Doe of his knowledge of the existence and location of potentially incriminating files; of his possession, control, and access to the encrypted portions of the drives; and of his capability to decrypt the files.
>
> We are unpersuaded by the Government's derivation of the key/combination analogy in arguing that Doe's production of the unencrypted files would be nothing more than a physical nontestimonial transfer. The Government attempts to avoid the analogy by arguing that it does not seek the combination or the key, but rather the contents. This argument badly misses the mark.[1]

1 670 F.3d 1335, 1346 (11th Cir. 2012).

8.58 The court was reinforced in this view because the government had failed to show that the purported testimony was a 'foregone conclusion'. It was not sufficient for the government to argue that the encrypted hard drives were capable of storing data, some of which might be incriminating. Tjoflat J noted that:

> … nothing in the record before us reveals that the Government knows whether any files existed and are located on the hard drives; what's more, nothing in the record illustrates that the Government knows with reasonable particularity that Doe is even capable of accessing the encrypted portions of the drives.[1]

1 670 F.3d 1335, 1346 (11th Cir. 2012).

8.59 The second ground for allowing the appeal was on the issue of immunity. The court concluded it was necessary to look beyond the act of production and determine what conduct the government had actually claimed to cover when granting immunity to Doe. Tjoflat J found that the District Court erred in limiting Doe's immunity to the U.S. government's use of his act of decryption and production while allowing the government derivative use of the evidence such act disclosed. In this instance, the judge concluded that the government had effectively declined to offer Doe constitutionally sufficient immunity.

8.60 *Doe* is clearly distinguishable from *Boucher* and *Fricosu*. In the latter two cases, the government was aware of what was on Boucher's computer because of his own actions in displaying them to the officers, and also the contents on Fricosu's laptop because a discussion she had about relevant matters over the telephone conversation with her ex-husband was recorded.[1]

1 For a more detailed discussion of this case, see Hanni Fakhoury, 'A combination or a key? The Fifth Amendment and privilege against compelled decryption' (2012) 9 Digital Evidence and Electronic Signature Law Review 81.

8.61 It has been pointed out that 'the use of encryption is one of the great, virtually insoluble dilemmas of cyberspace.'[1] Encryption provides for privacy, but also has the capacity to prevent law enforcement agencies from tackling criminals effectively. In this respect, since case law has provided imperfect guidance, it may be necessary to address this issue by way of legislation. Perhaps the most illustrative example of this was the litigation known as *Government's ex parte application for order compelling Apple Inc. to assist agents in search* before the US District Court of the Central District of California. The US Government seized an iPhone 5c believed to have belonged to Syed Rizwan Farook, an alleged terrorist who perpetrated an attack which killed 14 people and injured 22 others. The iPhone was protected by a passcode. Later generation iPhones have their contents encrypted by default, and the passcode acts as the password. Thus without the password it was not thought possible to obtain access to the device. It is also possible to set the iPhone to delete the contents if a set number of incorrect passcodes are entered. The US government was concerned that it would not be possible to obtain access to and find evidence on the iPhone with this encryption.

1 Phillip R Reitinger, 'Compelled production of plaintext and keys' (1996) 1 U Chi Legal F 171, 206; Greg S Sergienko, 'Self Incrimination and Cryptographic Keys' (2006) 2 Rich JL & Tech 1, [30], that 'producing a cryptographic key gives the document a testimonial content by decrypting the document and returning it into plaintext. Thus, the compulsory production of the key is the compulsory creation of testimonial content.' That the judicial authorities also disagree on this point is not surprising.

8.62 They sought an order under the All Writs Act requiring Apple Inc, the developers of the iPhone, to assist them in circumventing the encryption.[1] Contrary to what was reported in most media, the order did not require Apple to break the encryption, but rather Apple was ordered to provide reasonable technical assistance in bypassing or disabling the auto-erase function, enabling the FBI to submit passcodes to the device for electronic testing and ensuring that the device will not purposefully introduce any additional delay between passcode attempts.[2]

1 *Matter of the Search of an Apple iPhone Seized During the Execution of a Search Warrant on a Black Lexus IS300, California License Plate 35KGD203*, No. ED 15-0451M, 2016 WL 618401 (C.D. Cal. 16 February 2016).
2 *Matter of the Search of an Apple iPhone Seized During the Execution of a Search Warrant on a Black Lexus IS300, California License Plate 35KGD203*, No. ED 15-0451M (C.D. Cal. 16 February 2016), Order 2.

8.63 Apple firmly resisted the imposition of the order, arguing that to do so would hand unparalleled powers to the government, which would render meaningless data privacy laws. They argued that any process they put in place could be exploited by others, which meant that the privacy of all its customers would be put at risk.[1] Ultimately, the FBI were able to obtain access to the device with the help of an unnamed third-party, and therefore the order was not enforced.[2]

1 Tim Cook, 'A Message to Our Customers' (Apple, Inc, 16 February 2016) <www.apple.com/customer-letter/>.
2 Rob Crilly, 'FBI finds method to hack gunman's iPhone without Apple's help' *The Telegraph* (London, 29 March 2016) <www.telegraph.co.uk/technology/2016/03/29/fbi-finds-method-to-hack-gunmans-iphone-without-apples-help0/>.

8.64 In a similar case,[1] the government sought an order before a New York court requiring Apple, Inc. to bypass the passcode security on an Apple device on the basis that such an order would assist in the execution of a search warrant previously issued by the court. Orenstein J denied the government's motion, on the basis that the government had failed to establish that the All Writs Act permitted the relief it sought, partly because Congress has considered legislation that would achieve the same result, but had not adopted it. The judge also noted that a court, when deciding whether to take such discretionary action, was required to consider three additional factors:

> 1. the closeness of the relationship between the person or entity to whom the proposed writ is directed and the matter over which the court has jurisdiction;
>
> 2. the reasonableness of the burden to be imposed on the writ's subject; and
>
> 3. the necessity of the requested writ to aid the court's jurisdiction (which does replicate the second statutory element, despite the overlapping language).[2]

1 *In re Order requiring Apple, Inc, to assist in the execution of a search warrant issues by this Court*, 2015 WL 5920207; *In re Apple, Inc.*, 149 F.Supp.3d 341 (E.D.N.Y. 2016).
2 *In re Apple, Inc.*, 149 F.Supp.3d 341, 351 (E.D.N.Y. 2016).

8.65 Orenstein J said that even if the statute did apply, all three discretionary factors weighed against the issuing of the requested writ, and that the application would be denied as a matter of discretion, even if it is available as a matter of law.

8.66 The applications made and the reaction by Apple led to a major debate about encryption. Companies such as Apple and other mobile telephone operators are trying

to ensure that the user controls the encryption on their devices. They design the software to try and ensure that there is no 'backdoor' or flaws that they can be compelled to exploit. To a certain extent this is a simple business decision. If companies of devices can obtain access to users' material on their devices and are in fact compelled to do so, users who are interested in privacy will stop buying their products, thus harming the companies' revenue. Linked to this is the idea that a backdoor can be identified by anyone, which could mean that the device is considered unsafe, and sales will be reduced. More importantly, if the emphasis is placed on the user, the company will not be entangled in legal disputes involving the government. If the industry co-operates with the government, then it is seen to harm privacy, but if it resists, the allegations are that they are on the side of criminals. If the technology is wholly user-controlled, then the problem shifts away from the corporation.

8.67 If the technology cannot be circumvented and, as in some Federal circuits, the courts do not recognize the power to compel the disclosure of the password, then the information on the encrypted device is placed outside of the reach of most law enforcement personnel. However, the data remains fully accessible to the person holding the device – who can obtain access to its contents at will – but proof of his criminality will be in reality placed outside the reach of investigators.

8.68 Sensible legislation could strike the correct balance. Reitinger argues that 'permitting law enforcement to compel the production of keys when necessary, with judicial supervision as appropriate, is a minimal accommodation to the need for public security in a world in which criminals have an increasing array of sophisticated tools at their disposal.'[1] It is difficult to think of any other aspect of evidence where a suspect is allowed to wilfully hide evidence of his criminality from law enforcement and for this to be condoned by the criminal justice system. Using encryption, a person can hide thousands of paedophilic images on a device. They could obtain access to them every day but, if they took appropriate precautions,[2] law enforcement would find it almost impossible to prove that the offence has taken place.[3] That is not in the interests of society, and a new balance is required. This is a point made by Orenstein J in his concluding remarks:

> How best to balance those interests is a matter of critical importance to our society, and the need for an answer becomes more pressing daily, as the tide of technological advance flows ever farther past the boundaries of what seemed possible even a few decades ago. But that debate must happen today, and it must take place among legislators who are equipped to consider the technological and cultural realities of a world their predecessors could not begin to conceive. It would betray our constitutional heritage and our people's claim to democratic governance for a judge to pretend that our Founders already had that debate, and ended it, in 1789.
>
> Ultimately, the question to be answered in this matter, and in others like it across the country, is not whether the government should be able to force Apple to help it unlock a specific device; it is instead whether the All Writs Act resolves that issue and many others like it yet to come. (footnote omitted)[4]

1 Phillip R Reitinger, 'Compelled production of plaintext and keys' (1996) 1 U Chi Legal F 206, fn omitted.

2 Deleting caches, recent document lists, etc.

3 Keylogging software would only work if a single device was used to obtain access to the material (or the software would be required to be placed on each device) and if a regular Internet connection

was used. Covert surveillance (cameras) could be installed to show the material being accessed, but law enforcement would need to know which room the device was located in, and it could be difficult to obtain authorization to do so, depending on the level of intrusion this could cause (e.g. if it was on a tablet, it may be necessary to have devices in each room, which could be construed a gross invasion of privacy).

4 *In re Apple, Inc.*, 149 F.Supp.3d 341, 376 (E.D.N.Y. 2016).

The approach in Canada

8.69 Another perspective to this debate was added by the decision of the Canadian court in *R. v Beauchamp*.[1] In this case, an unusual application was brought. Rather than the law enforcement agency seeking access to encrypted data, the defence sought an order to require the Crown to disclose a copy of encrypted files located on a hard drive that had been seized by the police. The Crown had not been able to de-encrypt the files, and as a result had no knowledge of the data that was encrypted. It was agreed that the encrypted information was both potentially inculpatory and potentially exculpatory for the accused parties. The Crown submitted that the encrypted information was beyond its control, and although it was arguably in its possession, it was not in a format that the Crown was able to view it. The judge concluded that the Crown was in partial possession and control of the hard drives, but it had no knowledge of the information in the encrypted files. Smith J analysed the position as follows:

> The seizure by the police of the hard drives containing encrypted information is similar to the seizure of a locked safe which the police cannot open, containing documents which include both inculpatory and exculpatory evidence. The police or Crown would clearly be in possession or control of the safe, but if they did not have the key or combination and were unable to break the safe open, then they would not have knowledge of the contents of the safe. In this case, the Crown's control of the contents of the safe, which are known to one accused but not to the Crown, is not complete, as the Crown needs the key or combination, or in this case the password, in order to access the documents in the safe. The unique feature of this case is that the accused … has the key or password, which is necessary to complete the possession or control of the information in the safe.[2]

1 2008 Can LII 27481 (ON SC).

2 2008 Can LII 27481 (ON SC), [40].

8.70 For these reasons, the application for disclosure of a copy of the encrypted files in the hard drives was refused, although the judge indicated that the applicants could, at their option, obtain disclosure of the contents if they provided the password or key to the Crown and the Crown would then review the material. Had the application been allowed, it would have created an untenable situation. The state would have provided a file that only one party (the defence) could view. The defence would presumably extract the exculpatory evidence without giving the Crown sight of the inculpatory evidence. It is suggested that this decision struck the correct balance, which is to enable the defence to disclose the key so that both parties will have access to the plaintext material.

Concluding observations

8.71 The development of good quality encryption means that the user can legitimately encrypt his data to protect it. However, it is obvious that a criminal can also use encryption to hide his actions. The philosophical basis around the right to self-incrimination is of fundamental importance in any criminal justice system. The difficulty is in establishing a balance between the right not to incriminate oneself when accused by the state, and the rights of victims – often children – whose lives have been destroyed by men (for it is mostly men) who manipulate others for their own unsavoury, illegal and often unpleasant sexual pleasure, and the rights of the people in any society to be protected and safeguarded from the few who plan and undertake acts of mass murder. There is a fine but difficult balance to be made,[1] and in this chapter we describe how three different jurisdictions have approached the problem.

1 For a different perspective, see Phillip Rogaway, 'The moral character of cryptographic work', Cryptology ePrint Archive, Report 2015/1162, available at <http://web.cs.ucdavis.edu/~rogaway/papers/moral.html>.

9

Proof: the technical collection and examination of electronic evidence

Stephen Mason, Andrew Sheldon and Hein Dries

9.1 The activities associated with the investigation and examination of electronic evidence are relatively new compared to other forms of forensic analysis. A number of respected commentators who also practice as digital evidence professionals encourage their peers to advance the process of dealing with electronic evidence as a separate forensic science discipline.[1] This is reflected in the United Kingdom, where the government created a new post, that of the Forensic Science Regulator, in 2008. Under that post, a number of specialists groups were established, including the digital forensics specialist group. The Forensic Science Regulator is currently in the process of reviewing the broad range of standards and guidelines throughout the entire forensic industry, including that for digital forensics.[2] Broadly speaking, digital investigations are concerned with the gathering and analysis of relevant digital data to provide both evidence and intelligence to assist with an investigation in a criminal context.[3]

1 Fred Cohen, *Digital Forensic Evidence Examination* (4th edn, Fred Cohen & Associates 2012); Eoghan Casey, *Digital Evidence and Computer Crime: Forensic Science, Computers and the Internet* (3rd edn, Academic Press 2011), 1; Alastair Irons and Anastasia Konstadopoulou, 'Professionalism in digital forensics' (2007) 4 Digital Evidence and Electronic Signature Law Review 45; Simson Garfinkel, Paul Farrell, Vassil Roussev and George Dinolt, 'Bringing science to digital forensics with standardized forensic corpora' (2009) 6 Digital Investigation S2; Yinghua Guo, Jill Slay and Jason Beckett, 'Validation and verification of computer forensic software tools—Searching Function' (2009) 6 Digital Investigation S12; Simson L. Garfinkel 'Digital forensics research: The next 10 years' (2010) 7 Digital Investigation S64; Jason Beckett and Jill Slay, 'Scientific underpinnings and background to standards and accreditation in digital forensics' (2011) 8 Digital Investigation 114.

2 Codes of Practice and Conduct for forensic science providers and practitioners in the Criminal Justice System (Forensic Science Regulator, Version 3.0, February 2016); the Nederlands Register Gerechtelijk Deskundigen has begun a similar process (<https://lrgd.nl>).

3 The evidence of digital systems can also help reconstruct what happened in an incident, for which see Mario Piccinelli and Paolo Gubian, 'Modern ships Voyage Data Recorders: A forensics perspective on the Costa Concordia shipwreck' (2013) 10 Digital Investigation S41.

Guidelines for handling digital evidence

9.2 A number of guidelines, papers and other projects have been undertaken and published in relation to the collection and handling of electronic evidence,[1] and the digital forensic community has argued for a global response to the issue.[2] Those guidelines the authors are aware of are listed in Appendix 1.

1 Benjamin Turnbull, 'The adaptability of electronic evidence acquisition guides for new technologies' in *Proceedings of the 1st International Conference on Forensic Applications and Techniques in Telecommunications, Information, and Multimedia and Workshop* (ICST, Brussels 2008).

2 Part of (2011) 8 (2) Digital Investigation deals with standards, professionalization and quality in digital forensics; M Grobler, 'Digital Forensic Standards: International Progress' in Nathan Clarke, Steven Furnell and Rossouw Von Solms (eds), *Proceedings of the South African Information Security*

Stephen Mason, Andrew Sheldon and Hein Dries, 'Proof: the technical collection and examination of electronic evidence', in Stephen Mason and Daniel Seng (eds.), *Electronic Evidence* (4th edn, University of London 2017) 285–338.

Multi-Conference (University of Plymouth School Of Computing Communications and Electronics 2010).

9.3 International moves in relation to electronic evidence began in 1995, when the International Organization on Computer Evidence was established to provide international law enforcement agencies with a forum to facilitate the exchange of information relating to computer crime investigations and other issues relating to digital forensic investigations.[1] This organization, together with several other agencies, including the Association of Chief Police Officers and the National High-Tech Crime Unit, have produced a number of guidelines that influence the investigation and examination of electronic evidence within a criminal context. Although various sets of guidelines have, in the main, been produced specifically for criminal investigations, nevertheless, the guides that have been produced provide significant help to practitioners and lawyers in civil matters.[2]

1 See also N. Dudley-Gough, 'Digital forensic certification board' (2006) 3 Digital Investigation 7; Amber Schroader and N Dudley-Gough, 'The Institute of Computer Forensic Professionals' (2006) 3 Digital Investigation 9; note also the European Informatics Data Exchange Framework for Court and Evidence, a project running for 32 months (March 2014 – October 2016), <www.evidenceproject.eu>.
2 Casey, *Digital Evidence and Computer Crime* 230, indicates that the most mature and practical guidelines are those produced by ACPO.

Forensic triage

9.4 Preceding the investigation and examination of electronic evidence by digital evidence professionals is a new technique known as 'forensic triage',[1] which has received considerable attention within the forensic practitioner and law enforcement communities. Digital forensic triage is the term used to cover a range of processes, methodologies, software and hardware that can be used enable people to prioritise their digital forensic investigations more effectively. Forensic triage is not suitable for every case. It must be used in conjunction with appropriate risk assessment and by users with appropriate training. Indeed, there are direct comparisons to be drawn in this regard with the law enforcement and the medical profession. Applying the triage process, a police officer, trained in the use of a breath test meter, can use such a device to make informed decisions about a driver suspected of being intoxicated. The officer does not need to be an expert in the science embodied in the device but, instead, simply needs to be appropriately trained to configure, use and interpret the results it provides, and decide how best to take the investigation forward. Likewise, members of staff in accident and emergency departments are not all brain surgeons. Instead, some members of staff have appropriate training in using simple techniques and equipment that allow them to evaluate symptoms and direct the patient to the most appropriate treatment.

1 Marcus K. Rogers, James Goldman, Rick Mislan, Timothy Wedge and Steve Debrota, 'Computer forensics field triage process model' (2006) 1 Journal of Digital Forensics, Security and Law 19.

9.5 Various digital forensic triage methods, software, hardware and processes have been reviewed by the UK Centre for Applied Science and Technology (CAST). These evaluations, although focused on establishing if individual tools meet the claims made by the publishers, also test the effectiveness of the technology to preserve the integrity

of the target media to correctly identify specific digital artefacts and to produce results that would stand up to scrutiny using other forensic techniques. The outcomes of these independent tests were made available to police and other authorities under various classification restrictions, allowing them to form opinions about the suitability of each tool for given scenarios.

9.6 Digital forensic triage technologies and methods are still in their infancy,[1] and must take account of the need for appropriate training and accreditation. Similarly, suitable risk assessment will be required in order to minimize the omission of any relevant data. It could be argued that by not performing a full forensic examination of every piece of digital media found, vital evidence will be lost. Indeed, in some cases this may be true, but it should also be noted that, even when all devices submitted for examination have been scrutinized, it is still possible that not every device has been seized and its data assessed.

1 Faye Mitchell, 'The use of artificial intelligence in digital forensics: An introduction' (2010) 7 Digital Evidence and Electronic Signature Law Review 35.

9.7 An important consideration when employing digital triage techniques is the need to balance the rapid identification of material of interest and the consequence of stopping further analysis, knowing it is possible that such a process may fail to identify exculpatory material or material of more significance. By way of example, consider a hypothetical case of a user who is downloading indecent images of children from the Internet, and is also abusing his own child and uploading images that he has created. The investigator can use keyword and hash set analysis to quickly identify images known to be indecent on any media seized. If only the results of this triage examination method are presented to the accused at interview, the accused may make an early confession to charges of possession of the known images on the assumption that the investigation has failed to identify the more serious physical abuse and file sharing offences. Perhaps for this reason alone, digital triage techniques should be considered a powerful early investigation technique designed to enable investigators to make more informed decisions earlier in the media forensic processing cycle rather than being the only investigative technique used.

Handling electronic evidence

9.8 As with any other form of evidence, there are a number of discrete elements that accompany the collection and handling of digital evidence. It is suggested that a digital evidence professional should, ideally, undertake his duties against the highest standards that are propounded by their peers, regardless of whether he is advising in a criminal or civil matter. In *Bilta (UK) Limited (in Liquidation) v Nazir*,[1] Lewison J indicated that he did not consider it an automatic requirement that parties to civil proceedings have to subject hard drives to forensic discovery techniques. However, it is debatable whether it is wise not to subject hard drives to forensic discovery techniques, as demonstrated in the case of *In the matter of Stanford International Bank Limited (in liquidation), Fundora v Hamilton-Smith*.[2] This was an application for the removal of the Joint Official Liquidators of Stanford International Bank Limited, Nigel Hamilton-Smith and Peter Wastell, from their roles as Joint Official Liquidators, on the basis, amongst other reasons, that they destroyed digital data and employed improper

practices in relation to computer and electronic data.[3] The precise matters in dispute were as follows:

> The matters which tell [sic] to be considered can be narrowed down to the following: (a) three servers at the Montreal office of SIB were not imaged and not copied, (b) four desktops and laptops were not imaged but were securely erased, (c) the email servers and Blackberry enterprise servers were not imaged; (d) the IT specialists did not appear to have been instructed by the Liquidators to search for, collect and image the Blackberrys and data sticks.[4]

1 [2010] EWHC 3227 (CH), 2010 WL 4737753.

2 2-3 March and 8 June 2010, Claim Number ANUHCV2009/0149 Eastern Caribbean Supreme Court in the High Court of Justice Antigua and Barbuda; the judgment is available at <www.eccourts.org/wp-content/files_mf/1358795765_magicfields_pdf_file_upload_1_1.pdf> and the Court of Appeal decision is available at <www.eccourts.org/wp-content/files_mf/1358779099_magicfields_pdf_file_upload_1_1.pdf>.

3 Discussed at [44]–[115] of the judgment.

4 At [50] of the judgment.

9.9 In taking into account the relevant ACPO guidelines at the material time, Thomas J decided that the action of the Joint Official Liquidators was not in accordance with the standard forensic practice, and in so doing, they acted improperly.

9.10 To this extent, the various guidelines put forward as best practices provide sound advice and guidance when dealing with electronic evidence and can act, if followed, to counter allegations that the evidence has not been gathered or dealt with properly. This is because of the unique nature of electronic evidence: it is extremely volatile and subject to being altered with ease, even by the simple act of switching a computer on or off.[1]

1 Graeme B Ball and Richard Boddington, 'Solid state drives: The beginning of the end for current practice in digital forensic recovery?' (2010) 5 Journal of Digital Forensics, Security and Law 1 <http://ojs.jdfsl.org/index.php/jdfsl/article/viewFile/21/45>; Michael Wei, Laura M. Grupp, Frederick E Spada and Steven Swanson, 'Reliability erasing data from flash-based solid state drives', *Proceedings of the 9th USENIX Conference on File and Storage Technologies* (USENIX Association Berkeley, CA 2011).

9.11 In the case of *Khodorkovskiy and Lebedev v Russia*[1] before the European Court of Human Rights, the defence raised a number of important issues that relate to the volatile and mutable nature of electronic evidence:

> (i) A claim that the hard drives that were seized had not been properly packed and sealed, so it was possible to add information to them while the drives were in the possession of the General Prosecutor of the Russian Federation.[2]
>
> (ii) The hard drives seized during the searches of 9 October 2003 had been copied onto re-writable disks provided by the General Prosecutor, and were then transmitted to the experts without having been properly sealed. When examining the hard drives, the investigators discovered 4,939 more files than on the drives examined by the experts.[3]
>
> (iii) There were a number of failures relating to the data obtained from a server that was seized during the search in Zhukovka on 9 October 2003:
>
> > (a) the bill of indictment contained contradictory information on the location of servers, and on the type of the recording device where the information had been found;

(b) neither the hard drive nor the list of files discovered by the prosecution on it was attached by the General Prosecutor to the case materials;

(c) the files were copied by the experts to another hard disk, which had been provided by the General Prosecutor;

(d) the hard disk was re-writable, which meant it was possible to re-write and amend information on it – it could not be ascertained whether the hard disk had information on it before it was submitted to the experts;

(e) there was no evidence that documented the continuity of the evidence.[4]

(iv) When examining the hard disk, the investigators discovered more files than on the disk examined by the experts.[5]

1 11082/06 13772/05 – [2013] ECHR 747 (25 July 2013).
2 11082/06 13772/05 – [2013] ECHR 747 (25 July 2013), [72].
3 11082/06 13772/05 – [2013] ECHR 747 (25 July 2013), [181].
4 11082/06 13772/05 – [2013] ECHR 747 (25 July 2013), [678].
5 11082/06 13772/05 – [2013] ECHR 747 (25 July 2013), [679].

9.12 In concluding that these deficiencies were not relevant, the court said:

> Possible discrepancies in the documents describing the amount of data contained on the hard drives, inaccuracies as to the exact location of the computer servers, and other defects complained of may have various explanations. The Court cannot detect any manifest flaw in the process of seizing and examining the hard drives which would make the information obtained from them unfit for use at the trial.[1]

1 11082/06 13772/05 – [2013] ECHR 747 (25 July 2013), [702].

9.13 This is an extraordinary conclusion, given the importance attached by digital evidence professionals in ensuring strict adherence to proper guidelines to ensure the authenticity and integrity of data in digital form. The members of the court manifestly failed to understand the importance of digital forensics when seizing data in digital form, and have, arguably, undermined the importance of the issue in legal proceedings. Note that a hash value, calculated on site upon taking a forensic image of the seized hard drives (or, if this is impossible, shortly thereafter), could have easily served as proof of the evidence having been untouched since it was first acquired (provided that the hash was kept securely or communicated to the defence at an early stage). Yet there was no indication that the court or the defence understood this best practice.

Identifying electronic evidence

9.14 Evidence discovered in digital form may be the first sign that something is wrong. For instance, a security administrator in a bank might consider an investigation necessary when the intrusion detection system sets off an alarm, or where the email logs indicate that a particular member of staff is receiving an excessive number of emails during the course of a day or over an extended period. The case of *Miseroy v Barclays Bank plc*[1] illustrates the nature of the problems that are associated with the use of communication systems. Barclaycard employed Mr Hilary Miseroy in the Fraud Prevention Department between 14 March 1988 and 13 September 2002. The Staff Manual dated 16 June 2000 included a policy in relation to the supply and trafficking of drugs and money laundering. In addition, the Group IT Security Policies,

dated July 2002, included instructions about the use of the corporate email facilities. Barclays sent out clear guidance in both these areas. In July 2002, Maureen Crane, a Senior Fraud Analyst, was informed that an individual within her team appeared to be receiving a disproportionate number of emails during the day. A formal investigation was subsequently initiated. The Information, Risk and Security Department carried out an audit of the emails sent and received by three employees. The audit indicated that Mr Miseroy sent a significant number of emails. As a result, he was also included in the investigation. After a series of investigatory meetings, it was concluded that Mr Miseroy had abused the email facilities, as follows:

> (i) He sent out an unwarranted number of personal emails. On some days eight or more exchanges had taken place in quick succession.
>
> (ii) Some of the emails he sent out included content that was derogatory, offensive and sexist. During his first interview, he accepted that the comments he made were not appropriate. Later, he contended that there was a great deal of social activity and laddish banter between employees working within the Fraud Department and he did not consider that anybody had been offended.
>
> (iii) A number of emails were exchanged between him and Andrew West, a manager in a different department, between 26 April and 30 April 2002. The content of these emails referred to the purchase of cannabis from a friend of Mr Miseroy, who in turn passed the drug to Mr West. Similar emails had passed between Mr Miseroy and Mr West between 15 February and 10 April 2002. In an email dated 15 February, Mr Miseroy wrote to Mr West: 'I've brought it in with me. Fag-break about 10.30?' In a further email sent on 18 February, Mr Miseroy asked 'quality ok?'
>
> (iv) It was also determined that Mr Miseroy disclosed confidential information regarding Barclay's operations and customers.

1 (Case No 1201894/2002) (18 March 2003, unreported) Bedford employment tribunal.

9.15 Mr Miseroy was summarily dismissed for gross misconduct on 13 September 2002. The members of the tribunal accepted that the dismissal of Mr Miseroy was within the range of reasonable responses of a reasonable employer in relation to the circumstances of the case.

9.16 In such a case, the source and reliability of the information needs to be assessed, which requires an investigation into the facts. At such an early stage, the actions of the investigator may cause changes to the electronic evidence – for instance, in the case of *Aston Investments Limited v OJSC Russian Aluminium (Rusal)*, the actions of what appear to be the IT administrators were such that important files and information were removed, and subsequent forensic examination ran into difficulties because of the unintended changes made to the system.[1] This is why it is essential to have an appropriate procedure in place to deal with the way an investigation is initiated and conducted. In a civil case, there is an obligation for each party to disclose documents relating to matters in question in the action under the provisions of the Civil Procedure Rules.[2] In criminal matters, the relevant investigating authorities have both common law and statutory powers to search and seize evidence. In the criminal context, investigating police officers will be expected to have conducted themselves in accordance with the recognized guides for their jurisdiction. In the United Kingdom, ACPO has produced the *ACPO Good Practice Guide for Digital Evidence* (March 2012, v5) (ACPO Guide).[3] The ACPO Guide sets out four main phases for handling evidence:

collection, examination, analysis, and reporting, and concentrates on the collection phase. A digital evidence professional should consider adopting the same four phases for his investigations. With the advent of forensic triage techniques, these four phases may be augmented with an initial phase of 'assessment' or 'triage selection'.

1 [2006] EWHC 2545 (Comm).
2 For a discussion of some flaws in the legal and forensic process, see Vlasti Broucek, Paul Turner and Sandra Frings, 'Music piracy, universities and the Australian Federal Court: Issues for forensic computing specialists' (2005) 21 Computer L & Secur Rep 30.
3 Available at <http://library.college.police.uk/docs/acpo/digital-evidence-2012.pdf>.

9.17 While the following discussion concentrates on matters relating to electronic evidence in the context of a criminal investigation, the reader will readily acknowledge the relevance of the discussion in the context of a civil matter. As a result, a digital evidence professional, when undertaking work in the disclosure phase of a civil action, ought to be equally aware of the points that follow.[1]

1 The tension between forensics and investigations is discussed, amongst other things, in Monique Mattei Ferraro and Andrew Russell, 'Current issues confronting well-established computer-assisted child exploitation and computer crime task forces' (2004) 1 Digital Investigation 7.

Gathering electronic evidence

9.18 Once it has been established that it is necessary to seize or gather evidence in digital form, a further set of procedures should be in place to guide the digital evidence professional with respect to the scene itself, including the identification and seizure of the evidence if necessary.[1] It is now a well-established practice that the scene should be photographed, or even recorded by video, and the layout of the hardware recorded. The investigator then needs to determine what, if any, physical evidence, such as computers, printers, computer mice or facsimile machines, should be retained (the ACPO Guide provides a list of the types of hardware and storage devices that are susceptible to being retained).[2] It is important not to permit anybody to disturb the hardware or the network, or work on a computer that is liable to being seized and retained, and it is advisable that the police officers engaged in searching for digital evidence be properly trained.[3] The problem with digital evidence is the ease by which the data can be altered or destroyed. Digital devices are volatile instruments. For instance, the random access memory in a computer will contain a great deal of information relating to the state of the computer, such as the processes that are running, whether the computer is connected to the Internet, and what file systems are being used. When a computer is switched off, a large part of this volatile data is immediately and irretrievably lost. Depending on the circumstances of the case being investigated, it may be very important to retain such data before the computer is switched off or simply unplugged from the electricity supply. This question is becoming increasingly important because of the ready availability of encryption utilities that are easy to use, and the increasing availability of low cost hard disks that include whole disk encryption as a matter of course. The preservation of a forensic copy of a computer system's RAM may be the only way of gaining investigative access to the contents of a target device whose content is encrypted.[4] Indeed, there may be occasions when great care should be taken when arresting suspects caught physically at a computer, because it is possible that they might switch off the computer and disrupt or delete any incriminating files before any preventative action can be taken, as in the case of Aleksei Kostap. He was arrested

by members of the Serious and Organised Crime Agency, who attached handcuffs to him, but with his hands in front of his body. According to a press report, he managed to take action that caused certain databases to be deleted. It was thought the databases might have contained records of the gang's activities. Apparently, while handcuffed, Kostap also acted to initiate the use of intricate layers of encryption on the computer systems, which experts were not able to decrypt.[5]

1 For a brief discussion about gathering evidence and issues surrounding personal privacy, see María Verónica Péez Asinari, 'Legal constraints for the protection of privacy and personal data in electronic evidence handling', (2004) 18 International Review of Law, Computers & Technology 231.
2 Brian Carrier and Eugene H Spafford, 'Getting physical with the digital evidence process' (2003) 2 Intl J of Digital Evidence.
3 Although Harvey J in the District Court, Manakau in Canada, ruled that digital evidence was not necessarily rendered inadmissible because the accuracy of the data might have been jeopardised where a police officer, with full knowledge of the relevant guidelines, chose to ignore them. In this instance, during the search of premises a police officer switched on a computer and took 45 minutes to search various files stored on the computer: *R v Good* [2005] DCR 804. For problems when investigating mainframes and very large systems, see Matthew Pemble, 'Investigating around mainframes and other high-end systems: The revenge of big iron' (2004) 1 Digital Investigation 90.
4 Casey, *Digital Evidence and Computer Crime* 478.
5 Tom Espiner, 'Jailed ID thieves thwart cops with crypto', *ZEDNet UK* (19 December 2006).

9.19 In addition, new developments in the methods used to store data on storage devices may also cause problems in the future. Graeme B. Bell and Richard Boddington have demonstrated that:

> Evidence stored on modern internal primary storage devices can be subject to a process we label 'self-corrosion'. What is meant by this is that even in the absence of computer instructions, a modern solid-state storage device can permanently destroy evidence to a quite remarkable degree, during a short space of time, in a manner that a magnetic hard drive would not. Here, the phenomenon of solid-state drive (SSD) self-corrosion is proven to exist through experimentation using real world consumer hardware in an experimentally reproducible environment.[1]

1 'Solid state drives: The beginning of the end for current practice in digital forensic recovery?'

9.20 The authors provide a list of 21 recommendations, guidance and observations that digital evidence professionals may find of interest.

Copying electronic evidence

9.21 The process of copying (acquiring) and handling electronic evidence should be carried out to the highest standards, and is subject to several commonly applied best practices and principles. To this extent, the four principles of handling computer based electronic evidence as set out in the ACPO Guide[1] illustrate the importance of the data collection phase of this process:

> Principle 1: No action taken by law enforcement agencies, persons employed within those agencies or their agents should change data which may subsequently be relied upon in court.
>
> Principle 2: In circumstances where a person finds it necessary to access original data, that person must be competent to do so and be able to give evidence explaining the relevance and the implications of their actions.
>
> Principle 3: An audit trail or other record of all processes applied to digital

> evidence should be created and preserved. An independent third party should be able to examine those processes and achieve the same result.
>
> Principle 4: The person in charge of the investigation has overall responsibility for ensuring that the law and these principles are adhered to.

1 See Casey, *Digital Evidence and Computer Crime* 471 for a further discussion of documentation and a sample preservation form.

9.22 The problem of what to seize and retain can be compounded where a computer or an entire system of computers is linked to a network, and the sources of digital evidence exist in a number of separate geographical locations. In such circumstances, and before taking any action, it will be necessary to ascertain whether it is possible or feasible to shut the network down. In most instances, this will not be an option. Consequently the investigator will need to be aware of the range of original data that might be required, should they be presented with such a situation. This will include establishing the topology of the network that is to be investigated for the data, especially if a system administrator will not co-operate. For instance, it will probably be necessary to establish the number of computers on a network, and the various types of network connections such as Internet, cellular data networks and wireless connections available on the network.

9.23 Professor Casey posits two empirical laws of electronic evidence collection that ought to be high on the agenda:[1]

> Empirical Law of Digital Evidence Collection and Preservation 1: If you only make one copy of digital evidence, that evidence will be damaged or completely lost.
>
> Empirical Law of Digital Evidence Collection and Preservation 2: A forensic acquisition should contain at least the data that is accessible to a regular user of the computer.

1 Casey, *Digital Evidence and Computer Crime* 481.

9.24 To ensure a complete copy of a disk is obtained, Professor Casey recommends taking a bitstream copy of the electronic evidence.[1] As a result, the copy will include information that will normally enable a digital evidence professional to reconstruct deleted files, depending on the storage technology that was used. In circumstances where the volume of digital data is so large that copying it in its entirety is not possible,[2] it is generally accepted that 'logical' copies of the data may be made, provided the data that is copied can be shown to be an accurate and exact duplicate of the data that is the subject of copying, which can more usefully be described as 'first in time evidence'. Many methods exist for achieving this, including the use of proprietary 'logical evidence file formats' of common forensic tools.

1 Casey, *Digital Evidence and Computer Crime* 482.

2 This is reflected in the *Supplementary Attorney General's Guidelines On Disclosure: Digitally Stored Material* (14 July 2011), para 12.

9.25 There are two fundamental principles in relation to copying electronic evidence that a digital evidence professional should be aware of:

> (i) The process of making the image should not alter the first in time evidence. This means that appropriate steps should be taken to ensure that the process used to take the image should not write any data to the original medium.

> (ii) The process of copying data should produce an exact copy of the first in time evidence. Such a reproduction should allow the specialist to investigate the files in the way that they existed on the original medium.[1]

1 Troy Larson, 'The other side of civil discovery: disclosure and production of electronic records', in Casey (ed), *Handbook of Computer Crime Investigation*.

9.26 To ensure the first in time data and the copy are the same, the data should undergo a hashing process, described below. The reason for establishing hash values for data, including the time and date stamps of each file, is that this information will serve as a reference for checking the authenticity or veracity of the files after they have been copied.

9.27 The quality of digital files that are copied can be crucial. In the case of *The Gates Rubber Company v Bando Chemical Industries Limited*,[1] Schlatter MJ commented on the evidence of two digital evidence experts. The judge was impressed by the 'credentials, experience and knowledge' of the expert for *Bando*, Robert Wedig, and indicated in his decision that he relied on his opinions. As Gates failed to obtain an expert in a timely fashion, much less weight was placed on the expert for *Gates*.[2] The expert for *Gates*, Robert Voorhees, also failed to undertake appropriate measures to secure the evidence. Schlatter MJ's judgement is quoted more fully as follows to illustrate this point:

> Gates argued that Voorhees did an adequate job of copying the Denver computer. Wedig persuaded me, however, that Voorhees lost, or failed to capture, important information because of an inadequate effort. In using Norton's Unerase, Voorhees unnecessarily copied this program onto the Denver computer first, and thereby overwrote 7 to 8 percent of the hard drive before commencing his efforts to copy the contents.
>
> Wedig noted that information which is introduced into a computer is distributed, in a random manner, to space which is not being used, or to space which contains a deleted file and is therefore available for use. To use Norton's Unerase, it was unnecessary for Voorhees to copy it onto the hard drive of the Denver computer. By doing so, however, the program obliterated, at random, 7 to 8 percent of the information which would otherwise have been available. No one can ever know what items were overwritten by the Unerase program.
>
> Additionally, Voorhees did not obtain the creation dates of certain of the files which overwrote deleted files. This information would have assisted in determining the deletion date of some files. If a deleted file has been overwritten by a file which was created prior to the Gates litigation, for example, Bando would be relieved of suspicion as to that file. Thus, failure to obtain the creation dates of files represented a failure to preserve evidence which would have been important to Bando in its efforts to resist Gates' motions for default judgment.
>
> Wedig pointed out that Voorhees should have done an 'image backup' of the hard drive, which would have collected every piece of information on the hard drive, whether the information was allocated as a file or not. Instead, Voorhees did a 'file by file' backup, which copies only existing, nondeleted files on the hard drive. The technology for an image backup was available at the time of these events, though rarely used by anyone. Wedig testified that Gates was collecting evidence for judicial purposes; therefore, Gates had a duty to utilize the method which would yield the most complete and accurate results. I agree with Wedig. In these circumstances, Gates failed to preserve evidence in the most appropriate manner. Gates' failure to obtain an image backup of the computer is a factor which I have

weighed against Gates as I considered a number of the claims which Gates has asserted.[3]

1 167 F.R.D. 90 (D.Colo. 1996).
2 167 F.R.D. 90 (D.Colo. 1996), 111(a).
3 167 F.R.D. 90 (D.Colo. 1996), 112(a) and (b).

9.28 Although the tools and techniques used by digital evidence professionals are constantly changing and improving, nevertheless the comments made by the judge in this case illustrates a very clear point: when electronic evidence is copied, the techniques that are used ought to comply with the highest possible standards for the evidence to have any probative value in court, although it must be emphasized that there will be occasions when the investigator is faced with a unique situation such that she can only apply her knowledge to the best of her ability in seizing data in as forensic a way as possible. An example would be a live banking system. The system might be stored on hundreds of servers in a room the size of a football field, and the data will be changing every second. No set of guidelines cover such an eventuality, which is why the investigator must make decisions based on principles of good practice.[1]

1 For a sample imaging procedure, see Troy Larson, 'The other side of civil discovery: disclosure and production of electronic records', in Casey (ed), *Handbook of Computer Crime Investigation*; Barbara Guttman, James R Lyle and Richard Ayers, 'Ten years of computer forensic tool testing' (2011) 8 Digital Evidence and Electronic Signature Law Review 139.

9.29 An examination of the surrounding area of the scene, including any materials that are likely to be relevant to disclosure or a criminal investigation, is also important. For instance, in the case of *Regina v Pecciarich*[1] the police seized a number of documents, catalogues and a scrapbook of newspaper articles concerning trials of sexual assault and proposed legislation dealing with abusive images of children. In this instance, the material constituted real evidence. It was also considered, as Sparrow J determined, to be circumstantial evidence to support the allegations that Pecciarich distributed abusive images of children. The relevance of materials found at the scene, including fingerprints and DNA samples taken directly from hardware devices, may become more obvious once the digital evidence professional has examined the electronic evidence in detail.

1 22 OR (3d) 748.

Preserving electronic evidence

Validating digital data

9.30 Electronic evidence in particular needs to be validated if it is to have any probative value. A digital evidence professional will invariably copy the contents of a number of disks or storage devices, in both criminal and civil matters. To prove the electronic evidence has not been altered, it is necessary to put in place checks and balances to prove that the duplicate evidence in digital form has not been altered since it was copied. The method used to prove the integrity of data at the time the evidence was collected is known as electronic fingerprinting. The electronic fingerprint uses a cryptographic technique that is capable of being associated with a single file, a floppy disk or the entire contents of a hard drive. As electronic evidence is copied, so a digital evidence professional should use software tools that are relevant to the task.[1]

The software tool used will invariably incorporate a program that causes a checksum operation called a hash function to be applied to the file or disk that is being copied. When a hash function is applied to digital data, the result is called a hash value. The hash value has been calculated against the content of the data. This is a one-way function, containing the mathematical equivalent of a secret trapdoor. For the purposes of understanding the concept, this algorithm is easy to compute in one direction and difficult to compute in the opposite direction.[2] The hash function is used to verify that a file or the copy of a file has not changed. If the file has been altered in any way, the hash value will not be the same, and the investigator will be alerted to the discrepancy. However, it should be noted that Mr Luc Beirens, a Divisional Inspector of the Federal Police Service, Federal Computer Crime Unit, Belgian Federal Judicial Police, alerted those attending a judicial seminar in 2008 entitled *Investigation, Prosecution and Judgment of Information Technology Crime: Legal framework and criminal policy in the European Union,*[3] that he has experienced problems with some types of hardware, where certain sectors failed to function, which caused the hash algorithm to give different results than what was expected.

1 This is not what occurred in *State of Connecticut v Julie Amero* (Docket number CR-04-93292; Superior Court, New London Judicial District at Norwich, GA 21; 3, 4 and 5 January 2007) – for a detailed analysis of this case, see Stephen Mason (general editor), *International Electronic Evidence* (British Institute of International and Comparative Law 2008), xxxvi–lxxv; compare with the actions of the digital evidence professional David Hendricks in *Krause v State*, 243 S.W.3d 95 (Tex.App. 2007).

2 It has yet to be proven that a mathematical function can have a one-way function, see Fred Piper, Simon Blake-Wilson and John Mitchell, *Digital Signatures Security & Controls* (Information Systems Audit and Control Foundation 1999), 16.

3 This seminar was organized by the High Council of Justice, Belgium, in conjunction with the European Judicial Training Network, Ecole Nationale de la Magistrature (France), Consejo General del Poder Judicial (Escuela Judicial) (Spain), Studiecentrum Rechtspleging (Netherlands), Consiglio Superiore della Magistratura (Italy) and the Academy for Training of Judges and Prosecutors of the Republic of Macedonia, with financial support from the Directorate-General Justice, Freedom and Security of the European Commission (2007 Criminal Justice Programme) and the Federal Public Service Justice (Belgium), 25 November 2008 to 28 November 2008 at the Hôtel Jean de Bohême, Durbuy, Belgium. See also Mayank R. Gupta, Michael D. Hoeschele and Marcus K. Rogers, 'Hidden disk areas: HPA and DCO' (2006) 5 Intl J of Digital Evidence.

HASH collisions

9.31 There are many possible hashing algorithms that can be used to establish forensic veracity. For many years the MD5 (Message Digest 5) algorithm was used, but research conducted by Xiaoyun Wang and Hongbo Yuin showed that it was possible to create two files with different content that produced the same MD5 value.[1] The implications of this possibility quickly lead to some debate in the forensic community. One common interpretation was that MD5 could no longer be trusted because an analyst might wrongly identify an innocent file as a known file (the identification issue) or deliberately modify a file and change its hash value back to the original (the verification issue). Another hypothesis was that a suspect could make all his bad files have the hash values of known system files, thereby avoiding detection. While theoretically possible, it is practically very hard to achieve an MD5 hash collision and doing so requires serious computational time for files larger than a few hundred bytes. According to Stephens and others:

> It is important to note that the hash value shared by the two different files is a result of the collision construction process. We cannot target a given hash

> value, and produce a (meaningful) input bit string hashing to that given value. In cryptographic terms: our attack is an attack on collision resistance, not on preimage or second preimage resistance. This implies that both colliding files have to be specially prepared by the attacker …. Existing files with a known hash that have not been prepared in this way are not vulnerable.[2]

1 Xiaoyun Wang and Hongbo Yu, 'How to Break MD5 and Other Hash Functions', available at <http://merlot.usc.edu/csac-f06/papers/Wang05a.pdf>; Arjen Lenstra, Xiaoyun Wang and Benne de Weger, *Colliding X.509 Certificates* (version 1.0, 1st March 2005), available at <http://eprint.iacr.org/2005/067.pdf>; the earliest research is Hans Dobbertin, 'The Status of MD5 After a Recent Attack' (1996) 2 RSA Laboratories' CryptoBytes 1, 3–6.

2 Marc Stevens, Arjen K. Lenstra and Benne de Weger, 'Vulnerability of software integrity and code signing applications to chosen-prefix collisions for MD5' (30 November 2007), available at <www.win.tue.nl/hashclash/SoftIntCodeSign/>.

9.32 In mathematical terms, an MD5 hash is 128 bits wide and therefore the probability of two files having the same MD5 value is 2^{-128}. Put another way, the probability of finding two files with the same MD5 value is once in just over $3x10^{-39}$. That is one in 340 billion, billion, billion, billion comparisons. By contrast, a SHA-1 hash is 160 bits wide and so the probabilities increase to one in every 6.8×10^{-49} comparisons. In yet other words: in realistic terms, it is very hard, to the point of being practically impossible, to produce a 'doctored copy' of a larger digital evidence set that has the exact MD5 or SHA-1 hash value as the 'original' while still being 'believable'. However, this is not impossible, as the recent practical technique for generating an SHA-1 collision for PDF documents has demonstrated. It took the equivalent processing power of 6,500 years of single-CPU computations and 110 years of single-GPU computations, but resulted in a (believable) 'doctored copy' with a hash that was equal to a known original.[1] As SHA-1 is used extensively in the generation of trust certificates such as TSL and SSL, the publication of this new collision technique, together with proof of concept code to generate collisions, has resulted in at least one browser publisher (Firefox) removing SHA-1 as of 24 February 2017,[2] although Google Chrome had deprecated the use of SHA-1 since September 2014.[3]

1 Marc Stevens, Elie Bursztein, Pierre Karpman, Ange Albertini and Yarik Markov, 'The first collision for full SHA-1' (27 February 2017), available at <https://shattered.io/static/shattered.pdf>; John Leyden, Thomas Claburn and Chris Williams, ''First ever' SHA-1 hash collision calculated. All it took were five clever brains... and 6,610 years of processor time', *The Register* (23 February 2017).

2 <https://shattered.io>.

3 'Gradually sunsetting SHA-1', *Google* (5 September 2014), <https://security.googleblog.com/2014/09/gradually-sunsetting-sha-1.html>

9.33 The result of this debate is that, although the chance of an MD5 or SHA-1 collision is remote, the best practice suggests creating two hash values for every file or forensic image when used for comparison. If only a single hash algorithm is used, SHA256 would be better than MD5 or SHA-1. Using both MD5 and SHA-1 instead of a single SHA-256 is mathematically more robust. Further logic for this approach is the fact that although there are no national or international standards that require SHA-256 in digital forensics, its use instead of MD5/SHA-1 would immediately render all global child sexual exploitation image databases, which use MD5 and SHA-1 values, unusable. Furthermore, MD5 and SHA-1 are still used and accepted by every law enforcement authority worldwide to perform the three core forensic functions: to identify known indecent images, to exclude known files such as those in the National

Software Reference Library (NSRL) hash keeper list, and to verify that files have not been changed. In the light of the recent successful collision attack of SHA-1, this practice may need to be reviewed. It is advisable to retain first-in-time copies of any files that are to be identified, in order to be able to recalculate hashes as algorithms become deprecated and new ones are introduced. For digital signature purposes (detecting of changes) SHA-1 and MD5 should be considered unreliable and deprecated. Since digital signatures are usually only valid for a limited time period, this is less of a problem, although even with MD5, issues persist into 2017.[1]

1 Fahmida Y Rashid, 'Oracle to Java devs: Stop signing JAR files with MD5', *InfoWorld* (19 January 2017) <www.infoworld.com/article/3159186/security/oracle-to-java-devs-stop-signing-jar-files-with-md5.html>.

The continuity of custody

9.34 For those experienced in criminal matters, the concept of the continuity of custody (also known as the chain of evidence) is well established. However, the continuity of custody, in both civil and criminal matters, should be considered very carefully with respect to electronic evidence. The reason for taking particular care with electronic evidence is because it is easy to alter. It is necessary to demonstrate the integrity of the evidence and to show that it cannot have been tampered with after being seized or copied. There is another reason for being meticulous about ensuring the continuity of electronic evidence and that its custody is correctly recorded. In a case involving a number of items of hardware and more than one computer, it will be necessary to ensure that there is a clear link between the hardware and the electronic evidence copied from the hardware. In this respect, the record should address such issues as who collected the evidence, how and where it was collected, the name of the person who took possession of the evidence, how and where it was stored, the protection afforded to the evidence while in storage, and the names of the people who removed the evidence from storage, including the reasons for removing the evidence from storage.[1]

1 Warren G Kruse II and Jay G Heiser, *Computer Forensics Incident Response Essentials* (Addison-Wesley 2002), 6–11.

Transporting and storing electronic evidence

9.35 Consideration should be given to the methods by which any hardware and digital evidence is transported and stored.[1] Computers need to be protected from accidentally booting up, and consideration should be taken to ensure that hardware is clearly marked to prevent people from using the equipment unwittingly. Loose hard drives, modems, keyboards and other such materials should be placed in anti-static or aerated bags to prevent them from being damaged or their data corrupted. Storage conditions should be appropriate. Hardware and electronic evidence should be protected from dirt, humidity, fluids, extremes of temperature and strong magnetic fields. It is possible for data to be rendered unreadable if the storage media upon which the electronic evidence is contained are stored in a damp office or overheated vehicle. In many forensic storage facilities, special data safes protect evidence from fire risk. These safes are designed to withstand heat, and keep digital media at an acceptable temperature for longer periods of time during a fire.

1 Philip Turner, 'Unification of digital evidence from disparate sources (Digital Evidence Bag)' (2005) 2 Digital Investigation 223.

Cloud computing

9.36 In the same vein, evidence is increasingly stored with publically accessible, network based services. Both cloud computing (the use of, often shared or virtualized, computing or storage resources, available through the Internet) and the online delivery of services (software, infrastructure or platform as a service) are rapidly becoming more popular. Forensic investigation of these sources of evidence is inherently complex,[1] and is likely to force forensic standards involving the concept of 'original evidence' or 'first in time evidence' to become outdated or impracticable. In consequence, cloud forensics is emerging as a new and highly problematic field of computer forensics, disclosure, discovery and criminal investigations.[2]

1 Eoghan Casey, 'Cloud computing and digital forensics' (2012) 9 Digital Investigation 69; M Taylor, J Haggerty, D Gresty and R Hegarty, 'Digital evidence in cloud computing systems' (2010) 26 Computer Law & Security Review 304.

2 Stephen Mason and Esther George, 'Digital evidence and "cloud" computing' (2011) 27 Computer Law & Security Review 524; Ian Walden, *Law Enforcement Access in a Cloud Environment*, Legal Studies Research Paper No 74/2011, available at <http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1781067>; Giuseppe Vaciago, 'Remote forensics and cloud computing: An Italian and European legal overview', (2011) 8 Digital Evidence and Electronic Signature Law Review 124.

9.37 At the same time, the use of free or low cost 'cloud storage' adds a challenging complication to the process of preserving digital evidence. At the heart of this technology is the concept that a user can upload and store data and software applications to 'the cloud', which can then be accessed from anywhere using any device with an Internet connection. In reality, such 'cloud storage' can consist of many hundreds (or thousands) of mass storage devices (arrays of high capacity hard disks) located in many different physical locations, all connected to a storage management system software via the Internet. It is often the case that, in order to ensure that users' data is available at all times and to protect them from loss such as disk failure or interruptions in network connectivity, many copies of the users' data are spread across many redundant storage nodes, physically and geographically separated from one another. Furthermore, many unrelated users share cloud storage facilities. In these 'multi-tenanted' systems, the management of such data is largely automatic and controlled by the storage management system software rather than human managers. The implications for forensic preservation of such data are readily apparent. Approaching the operators of a 'cloud storage' product in order to gain access to a user's data may prove fruitless. This is because it is probable that not only will the operators not personally know where the data is stored, geographically or physically, but the data will almost certainly be automatically encrypted using a user-generated key that will not be known to the operators.[1] It follows that because of the geographically distributed nature of such systems, issues of legal jurisdiction may arise when seeking to preserve or obtain the data with the co-operation of the cloud operators.

1 For a general overview of some of the issues, see the entire issue of (2011) 14 (1) IAnewsletter, entitled 'Cyber forensics in the cloud' <www.csiac.org/wp-content/uploads/2016/02/Vol14_No1.pdf>.

9.38 One method of securing access to such data is to request that the user provides details of his account to enable suitably authorized police officers to log into the relevant account and copy all pertinent data to a forensic store disk or, more efficiently, copied from the user's 'cloud' to a storage location on a 'forensic cloud'. The forensic process should include the creation of hash values (discussed below) for every file and the use of automatic or manual logging of each action to create a contemporaneous note for all actions undertaken. Additionally it may be prudent, with the appropriate legal authorization, to change the access credentials of the original 'cloud' storage in order to prevent any deliberate or inadvertent changes from being made to it. In such circumstances, the provisions of ACPO principles 2 and 3 should be considered:

> Principle 2: In circumstances where a person finds it necessary to access original data, that person must be competent to do so and be able to give evidence explaining the relevance and the implications of their actions.
>
> Principle 3: An audit trail or other record of all processes applied to digital evidence should be created and preserved. An independent third party should be able to examine those processes and achieve the same result.

9.39 Furthermore, suitable additional steps must be taken such that all reasonable precautions to ensure that any changes are kept to a minimum, the changes are noted and recorded and the person conducting the acquisition process is fully aware of the effect of her actions. For those occasions when permission to obtain access to the data from the accused is not forthcoming, it is then essential, if possible, to preserve a copy of the volatile memory in the computer, if it is seized, so that it is possible to search for any remaining data relating to the account.

9.40 Data can be deleted on a remote server or cloud storage before it can be secured.[1] In such complex scenarios as described above, the role of forensic triage becomes increasingly important, because it allows the responding investigators to evaluate the scene contemporaneously, and to potentially identify the data, seek the appropriate authority to search and seize the data (if such an order or warrant has not been obtained, or if the order or warrant under which the search is being conducted does not cover the materials that have been found), and secure the online data before anybody who might be under suspicion (or their accomplices) gets the opportunity to destroy it remotely. It is in such circumstances that conducting a preliminary risk assessment is essential to success.

1 For a discussion of the complexities of recovering data from modern operating systems and file systems, see Geoff H Fellows, 'The joys of complexity and the deleted file' (2005) 2 Digital Investigation 89.

9.41 Throughout this phase of any investigation, the emphasis will be on the digital evidence professional to make informed decisions as to what equipment to seize and retain in any given set of circumstances.[1] Depending on the circumstances of the case, consideration has to be given to the possibility that the person at the centre of the investigation might be framed.[2] It will also be necessary to give reasons for seizing and retaining the property, and it will be essential to ensure that the entire procedure is properly documented. The documentation relating to electronic evidence is important. Standard operating procedures such as those described in the ACPO Guide, as noted above, should be followed. A record should be kept of every item seized, every action

performed that may affect electronic data on every item, and exhibit labels should be attached to every physical item retrieved.

1 The prosecution failed to analyse the evidence from the family computer effectively in the case of the death of Casey Marie Anthony in 2011, for which see Craig Wilson, 'Digital Evidence Discrepancies – Casey Anthony Trial' (11 July 2011) <www.digital-detective.net/digital-evidence-discrepancies-casey-anthony-trial/>; Tony Pipitone, 'Cops, prosecutors botched Casey Anthony evidence' (Clickorlando.com, 28 November 2012) <www.clickorlando.com/news/cops-prosecutors-botched-casey-anthony-evidence>; Jose Baez and Peter Golenbock, *Presumed Guilty: Casey Anthony: The Inside Story* (updated edn, BenBella Books 2013) 46, 180–3, 211, 346–8, 365, 368–71, 400, 426–8; Jess Ashton and Lisa Pulitzer, *Imperfect Justice: Prosecuting Casey Anthony* (William Morrow 2011) 105, 239, 277, 291–2, 298, 315.
2 John Leyden, 'Child abuse frame-up backfires on stalker', *The Register* (6 April 2010), in which Ilkka Karttunen broke into the Essex home of a woman he wanted to be with, downloaded abusive images of children on to the computer, then stole the hard drive and sent it into the police with a note identifying the owner; for a similar example, see 'Handyman jailed for planting porn on boss's computer', *BBC News London* (23 September 2010).

9.42 There are occasions when the physical hardware cannot be seized, because it is too large, it is not physically located in the jurisdiction, or where seizing it would cause an organization to cease functioning. In such circumstances, the digital evidence will have to be copied exactly. As a result, greater care must be exercised when electronic evidence is retrieved and copied for the first time. The range of electronic evidence that might need to be copied includes audit trails, data logs (for applications, Internet access[1] and firewall traffic, to name a few), biometric data, metadata from applications, file systems,[2] intrusion detection reports and contents of databases and files. Given the nature of the evidence to be copied, the evidential continuity and integrity of the evidence that is copied and its subsequent history become paramount.[3]

1 For an interesting discussion, see Richard Clayton, *Online traceability: who did that? Technical expert report on collecting robust evidence of copyright infringement through peer-to-peer filesharing* (Consumer Focus 2012) <www.cl.cam.ac.uk/~rnc1/Online-traceability.pdf>.
2 Florian Buchholz and Eugene Spafford, 'On the role of file system metadata in digital forensics' (2004) 1 Digital Investigation 298.
3 The volume of digital evidence is causing problems respecting the methodologies around the collection of evidence, as discussed in the US context by Erin E Kenneally and Christopher L T Brown, 'Risk sensitive digital evidence collection' (2005) 2 Digital Investigation 101; Simon Attfield and Ann Blandford, 'E-disclosure viewed as "sensemaking" with computers: The challenge of "frames"' (2008) 5 Digital Evidence and Electronic Signature Law Review 62; Daniel R Rizzolo, 'Legal privilege and the high cost of electronic discovery in the United States: Should we be thinking like lawyers?' (2009) 6 Digital Evidence and Electronic Signature Law Review 139.

9.43 Another way of dealing with this challenge is to request the co-operation of the service provider to retrieve forensic evidence copies of the information on its systems. This, however, may well lead to issues of jurisdiction. Thus the need for better guidance on the issues arising out of cloud computing is becoming clear. The Council of Europe has established a working group to address the issue and explore solutions on criminal justice access to evidence stored on servers in the cloud and in foreign jurisdictions, including through the process of mutual legal assistance.[1]

1 <www.coe.int/en/web/cybercrime/ceg>.

Analysis of electronic evidence

9.44 A digital evidence professional is not only required to obtain and copy electronic evidence that has a high probative value, but must also provide an analysis of the evidence. The analysis of the evidence will involve reviewing the text of the data, and the attributes of the data. This exercise may also include, but will not be limited to, looking for and recovering deleted files and other data that may be hidden on the disk, checking logs for activity and checking unallocated and slack space for residual data. Failure to assess the digital evidence can lead to false assumptions, as in the case of *Liser v Smith*.[1] Jason Liser was arrested for the murder of Vidalina Semino Door after being identified as a person who withdrew money from an automatic teller machine (ATM) owned by the Bank of America on the night of the murder. The facts of the case were not in dispute. The victim was shot after leaving work on the night of 5 May 2000. Police attended the scene shortly after the murder. By Monday 8 May, it was known that the victim's bank card had been used to withdraw US$200 from a Bank of America Branch about 20 minutes after the murder, approximately one mile from where the body was found. According to the electronic evidence, the withdrawal occurred at 1.47am on 6 May. A further US$81 was taken out of another ATM owned by a 7-11 store at 2.17am. The Bank of America ATM also had a video surveillance tape, which was subsequently retrieved by the police. The ATM at the 7-11 store did not have a working video camera.

1 254 F.Supp.2d 89 (D.D.C. 2003).

9.45 The bank manager informed the police that there would be a discrepancy of up to 15 minutes between the time indicated on the surveillance tape and the actual time. When the tape was viewed, there was no ATM activity recorded at 1.47am. The closest transaction that occurred was at 1.52am, when a black male wearing a white t-shirt (Jason Liser) was recorded as standing before the machine. While the evidence seemed to lead to the conclusion that the man recorded at 1.52am was one of the killers, the evidence contained on the surveillance video did not warrant such an assumption. Other pictures from the videotape showed black males other than Liser using the ATM at 1.56am and 2.05am, together with a black female using the machine at 2.04am. Copies of these pictures were provided to the court. All of them were grainy and poorly photocopied. However, of relevance was that both of the men in question appeared, like Liser, to have been wearing white t-shirts and to be relatively young.

9.46 In August 2000, it was decided to send out a press release and a copy of the photograph of the man recorded as standing at the ATM at 1.52am. Mr Liser was subsequently recognized and arrested. He was held for less than a week, because the police decided, at this late point in time, to carry out an experiment at the Anacostia branch. The result of the experiment led the police to conclude that the discrepancy was greater than the 15 minute gap they were led to believe. Mr Liser was subsequently released. It is instructive to note the comments made in the Memorandum Opinion by the judge:

> While this issue is a close one, the Court is not ready to conclude that it was objectively reasonable under the circumstances of this investigation for the police to rely solely on the bank's representations about the time discrepancy without attempting to verify that information by empirical (or other) means. The

> crucial point here is that this was not a fast moving investigation in which the officers were called upon to make snap judgments based on limited information. Far from it. Detective Smith had the surveillance tapes within a week after the murder; at that early date he had been told by the branch manager that the time on the tape could be off by up to fifteen minutes. … Plaintiff was not, however, arrested until August, three months later. During this lengthy interval, neither Detective Smith nor anyone on his team made any further attempt to verify the estimation about the length of the gap. They had no further contact with anyone at Bank of America, especially its security personnel, who might have had more accurate information about the camera's timer. … They did not inspect the camera itself. Nor did they attempt [to] use the ATM themselves to compare real time against tape time.
>
> In short, despite the fact that the tape was their central lead as to the identity of the murderer, the investigators did nothing to pin down exactly how far off the video clock was, at least not before plaintiff was arrested. [Footnote 3: The fact that the police finally sought to verify the information – and quickly and readily learned that it was inaccurate – *after* Liser's arrest certainly does not help their cause. That such an [sic] simple test was not done in the three months preceding the arrest, and if done would have cast serious doubt on the propriety of that arrest, suggests an investigative sloppiness that at least casts doubt on whether the initial arrest was actually supported by probable cause.] Instead, Detective Smith and his team chose to rely solely on a single, untested statement from the bank manager. Such reliance might well have been unassailable had the investigators been making an on-the-spot determination as to whether probable cause existed to arrest plaintiff in the first frantic days after the murder. But in the circumstances of the deliberate, slowly unfolding investigation that ensued, during which the officers should have had ample time to pursue leads and to check facts, their failure to verify the length of the gap on the video stands in a rather different light. Their conduct appears more sloppy than reasoned, the product of carelessness rather than craft. The Court is thus unable to say with certainty that this crucial mistake was ultimately a permissible one, or that prudent investigators would necessarily have conducted themselves as defendants did here.[1]

1 United States District Court for the District of Columbia No 00-2325 (ESH) 26 March 2003 before Ellen Segal Huvelle DJ, 11–12.

9.47 Compare this case with the murder of Denise Mansfield, who was found bound and strangled in her home on 29 June 2002. It was thought that she had been dead since 22 June. The police investigation centred on a surveillance camera that recorded images of people using an ATM, owned by the Sun Trust Bank. This ATM was used to withdraw US$200 from the victim's bank account at 2.30pm on 22 June, using her debit card. Three women (Virginia Shelton, her daughter Shirley and one of her daughter's friends, Jennifer Starkey) were subsequently arrested. They were identified as using the machine between 2.28pm and 2.33pm the same day. The women did not dispute using this particular ATM. They were subsequently released after three weeks. After they were arrested, it came to light that it was assumed the clocks on the transaction computer and the ATM were synchronized. This was not correct. The women used the ATM earlier than the time stamp on the video recording. It was reported that police officers had these records in their possession on the day they arrested the women, but it was not clear if they had examined the records before making the arrests. It was not until Mr Starkey obtained a copy of the relevant records that the women were released.[1]

1 Ruben Castaneda, 'Mistaken arrests leave Pr. George's murder unsolved' (washingtonpost.com, 22 June 2003) <www.washingtonpost.com/archive/politics/2003/06/22/mistaken-arrests-leave-pr-georges-murder-unsolved/8e6257de-22c6-4e73-894f-0e71f7ad9b2c/>. Previous online references to this article now omit the text of the article, but it remains available in full at <www.truthinjustice.org/PGfalse-confessions.htm>.

9.48 The case of *Liser v Smith* is a good example of the failure to fully test the electronic evidence, in particular, the time. No clock is accurate. This can be important in terms of assessing evidence in digital form.[1] In the legal context, Lord Hoffman observed, in *DPP v McKeown; DPP v Jones* that 'The clock, although no doubt physically in the same box as the computer, is something which supplies information to the computer rather than being part of the processing mechanism'.[2] It might have been correct that the clock was one hour out because of the difference in time zones, but clocks in computers are not always accurate. Clocks on facsimile machines may also be far from accurate, but the following comments by Burton J (President) in *Woodward v Abbey National plc* that imply that the data recorded by the logs at the offices of the Employment Appeals Tribunal are accurate as a matter of 'common sense' cannot be correct:

> [I]t must make common sense to accept the accuracy, as I believe there to be, of the record of receipt in the fax log of the [Employment Appeals Tribunal (EAT)], and not to accept either uncertain evidence about the accuracy of the sender's machine or some kind of speculation as to electronic receipt short of the record in the EAT fax log.[3]

1 The first voice in the play by Dylan Thomas, *Under Milk Wood*, referred to 'slow clocks, quick clocks' at [60], and the narrator in *The Time Regulation Institute* by Ahmet Hamdi Tanpınar (Maureen Freely and Alexander Dawe tr, Penguin 2014), [11] tells the reader that 'Everyone knows that a watch or clock is either fast or slow. For timepieces, there is no third state'. Dr John C Taylor invented, designed and gave the Corpus Chronophage to Corpus Christi College in Cambridge, England. It is a mechanical clock designed to demonstrate the principle of relative time, doing the unexpected, and is only accurate once every five minutes. The Chief Scientist for Time Services at the US Naval Observatory, Dr Demetrios Matsakis, is responsible for precise time determination and the management of time dissemination. To achieve this, there is a USNO Master Clock that is in turn based on a system of a number of independently operating cesium atomic clocks and hydrogen maser clocks, all of which automatically compare with each other, so that rate does not change by more than about 100 picoseconds (0.000 000 000 1 seconds) per day from day to day: <http://tycho.usno.navy.mil/clocks.html>.

2 [1997] 1 All ER 737, 754d; [1997] 1WLR 295; [1997] 2 Cr App R 155, HL.

3 *Woodward v Abbey National plc; J P Garrett Electrical Limited v Cotton* [2005] ICR 1702, [2005] IRLR 782, [14]. See his further comment on both cases in *Woodward v Abbey National plc; J P Garrett Electrical Limited v Cotton* (26 July 2005, unreported) (UKEATPA/0534/05/SM and UKEATPA/0030/05/DZM), and similar comments on the same point in *Midland Packaging Limited v Clark* [2005] 2 AER 266 EATPA/1146/04. In *R v Good* [2005] DCR 804 the clock in the computer was running 42 minutes and 30 seconds behind the actual time.

9.49 A more realistic comment on the accuracy or otherwise of clocks was made by Smart AJ in the case of *R v Ross Magoulias*,[1] where the identity of the appellant centred on the recordings made by an ATM and a security video:

> It is a notorious matter of fact that reliable clocks or timing devices may show slightly different times. A clock may gain or lose ever so slightly and it may be some days before the difference becomes noticeable. When setting a clock or timing device there might be a very small error. Perhaps the clock from which the timing device is set is slightly astray. It is exceedingly well known that the timing of differing clocks needs to be synchronised if pinpoint accuracy is required. It is beyond argument that both [the victim] and the appellant attended the service

> station on 7 July 2001. She can be seen on the video tape for about three minutes (18.37.18 to 18.40.25 according to the video tape timing device). That cannot be disputed. Nor can it be disputed that the appellant attended at the ATM and withdrew $50 (18.40.59 according to the ATM timing device). As earlier pointed out there was no direct evidence available to the jury that the timing mechanisms were not synchronised. If there had been the video tape would have recorded a person (the appellant) withdrawing $50 from the appellant's account at 18.40.59 (bank record time). The video does not show anybody near the ATM at that time. Thus there was no room for any presumption to operate in any useful way.

1 [2003] NSWCCA 143, [41].

9.50 A clock can help reveal the truth when somebody attempts to alter digital evidence, as in the case of Shaun Richards, a driving instructor, who falsified data from his satellite navigation system in an attempt to evade a speeding charge. He forgot to change the time on the product back to British Summer Time from Greenwich Mean Time, which meant that the clock was out by one hour. He was imprisoned for four months.[1]

1 'Devon driving instructor jailed for sat-nav speed fraud', *BBC News Devon* (13 January 2011).

9.51 Nevertheless there may be occasions where, in the absence of proof, an intelligent assumption that comments recorded on a document have a certain meaning might be accepted by an adjudicator, even when it is possible that the comments are capable of other meanings. In particular, the failure to offer an explanation to rebut the assumed meaning of the content of a digital document submitted in evidence may lead to a finding against the party adducing the evidence, as in *Hedrich v Standard Bank London Limited*.[1] The case concerned a wasted costs order, which was based on breach of the duty owed by a solicitor to the court to perform his duty as an officer of the court in promoting the cause of justice. Ward LJ took particular care in assessing the conflicting evidence, because of the complexity of the facts. The bank sought to have its costs paid by the claimants' solicitors, Messrs Zimmers. The bank was required to establish a strong prima facie case to succeed, and as part of their case, the bank sought to prove Zimmers were in receipt of an email on a date before Zimmers claimed that they had actual sight of the evidence. The bank relied on the truth of the content of the email. The relevant text that it relied on read:

> No virus found in this incoming message.
>
> Checked by AVG Free Edition.
>
> Version: 7.1.362/Virus Database: 267.12.8/162-Release Date: 05/11/2005.[2]

1 [2008] EWCA Civ 915.
2 [2008] EWCA Civ 915, [70].

9.52 In the absence of evidence from a digital evidence professional, the inference the bank sought to draw from this information was that the solicitors received notification of this particular email in May 2005, to counter the claim that they did not see it until the trial was under way in December 2005. This was highly relevant, because the bank was asking the court to order Zimmers to pay costs of £342,917.08. In meeting this argument, the barrister for Zimmers, Graeme McPherson QC, conducted some research on the Internet for an alternative explanation as to why the date was printed as 11 May 2005. Ward LJ accepted the following explanation offered, although there was no evidence of the truth of the alternative explanation:

> Mr McPherson's researches on the internet gave him an alternative explanation. He told us that the first line showed, as it states, that no virus had been detected. The second line indicates that the means of checking was by the AVG Free Edition, which is a free virus detection software programme marketed as AVG. The third line identifies the version of AVG's software and the crucial date upon which the Bank relies is simply, as is stated on the e-mail, the date of the release of that particular version of the software. We have no evidence that this is the true explanation: we only have Mr McPherson's word that his researches on the internet produced that answer. It may have been a moment of inspiration by counsel but for my part it has a compelling ring of truth and I have no reason to think that it is unreliable. It destroys that part of the Bank's case.[1]

1 [2008] EWCA Civ 915, [71].

9.53 It would have been wise of the bank to establish the meaning of this information, because of the evidential hurdle required to prove its case. It would not have taken a digital evidence professional long to have established whether the information proved the date was the date of the release of that particular version of the software or not. It might have been for the court to ask the parties to seek an opinion on this issue before reaching a conclusion, but given the nature of the proceedings, in particular the rule that where there is room for doubt, the respondent lawyers are entitled to the benefit of it, it is not surprising that the court did not let the matter continue any further, and accepted the alternative explanation.[1]

1 See the analysis of *State of Connecticut v Julie Amero*, Docket number CR-04-93292, Superior Court, New London Judicial District at Norwich, GA 21 that has a similar point, but the digital evidence professional for the prosecution failed to even consider looking for malicious software: Mason, *International Electronic Evidence*, xxxvi–lxxv.

9.54 A further observation of relevance is that in itself, the electronic evidence may not be conclusive. The case of *Mogford v Secretary of State for Education and Skills*[1] illustrates this point. Mr Mogford appealed against a decision of the Secretary of State for Education and Skills to include his name in the list maintained under the provisions of the Education (Restriction of Employment) Regulations 2000 (SI 2000/2419) that prevented him from being employed as a teacher under the provisions of regulation 5(1)(c). The Secretary of State made this decision because abusive images of children, text files, emails relating to this material, and bookmarks with links to websites containing abusive images of children had been found on Mr Mogford's computer. He denied that he was responsible for this material. The members of the Tribunal were satisfied that the Secretary of State proved on a balance of probabilities that either Mr Mogford was solely responsible for the materials found on the computer, or that he participated with others in obtaining this material, and he knew that it was on his machine. The reasons given included:

> (i) Inconsistencies in Mr Mogford's evidence. He frequently changed his story. He told the interview team that he was visiting his girlfriend on the weekend 25–27 April 1997, then changed his story before the members of the Tribunal, indicating that three people had stayed at his house that weekend. Another inconsistency relates to exactly who set up the Internet on his computer. He said in the interview that RS had helped set up his Internet link. In evidence to the Tribunal, RS indicated that this was not correct. Mr Mogford gave evidence to the effect that P set up the Internet for him.

> (ii) There was no attempt to find P, or indeed either of the other two friends whom Mr Mogford claimed were with him that weekend. That he failed to take steps to ask his friends to corroborate his story was held by the members of the Tribunal as being consistent with the fact that his version of events was not credible.

1 [2002] EWCST 11(PC) (26 June 2002).

9.55 Consideration was also given to the timing of the file system activity, and the members of the Tribunal carefully examined the evidence presented by the digital evidence professional, Mr T. Mr Mogford had created a spreadsheet that contained details of earnings from private lessons, and this spreadsheet was closed down at 00.28 on 27 April 1997. Mr Mogford claimed that he had opened his spreadsheet at some other time earlier, and failed to close it down. The members of the Tribunal articulated the importance of this item of evidence and the explanation offered by Mr Mogford as follows:

> It is our interpretation of the evidence that Mr M must have been using the computer at this time, either alone or with someone else, surfing the net and finding child pornography sites and text messages, and therefore when closing down the computer his spreadsheet would have been closed. The spreadsheet would have been of no interest to his friends, and he himself said in evidence that it was unlikely that he would have opened the spreadsheet and left it for a couple of days. We can only infer that he was working on the spreadsheet earlier that evening or the previous day.[1]

1 [2002] EWCST 11(PC) (26 June 2002), [25].

9.56 The observations noted above illustrate the importance of understanding the nature of digital data, as noted by the Legal Affairs Expert Panel of the British Computer Society that submitted comments to the *Criminal Courts Review* by Lord Justice Auld:

> A universal source of delay and wasted resource is the confusion shown by witnesses and lawyers between fact, conjecture, speculation, assumption, inference and opinion on technical matters. This phenomenon is usually closely related to the reluctance to consider multiple interpretations described above.
>
> The most common example is the confusion shown by technical witnesses and lawyers over the precise significance, and reliability as evidence, of the file date- and time-stamps recorded by a computer.[1]

1 <www.computerevidence.co.uk/Papers/LJAuld/BCSComputerEvidenceSubmission.htm>.

9.57 The aim should be to test the accuracy of the evidence and to ask if the conclusions are correct, rather than making decisions based on an imperfect analysis of the available evidence. It should never be assumed that because evidence is in electronic form, that it must therefore be correct and impervious to being tested to prove whether it is accurate or false.

Tools

9.58 A digital evidence professional will not only, ideally, require an in-depth knowledge of the operating system she is to investigate, but she will also need to use a number of proprietary tools in the performance of her investigation and analysis of digital evidence. The types of tool she uses will depend on the operating system she

is to look at, and whether she is investigating networks, handheld devices, embedded systems or wireless networks.[1] The specialist digital evidence textbooks consider these matters in depth, and the reader is encouraged to familiarise himself with the technology and techniques by referring to appropriate practitioner texts,[2] including their limitations.[3] The tools used can, naturally, be the subject of cross-examination, and the underlying scientific methodology and structure of such tools can also be questioned.[4] In this section, the aim is to illustrate why and how tools are used in the context of the Windows operating system, partly because it is so popular.

1 W Jansen and R Ayers, 'An overview and analysis of PDA forensic tools' (2005) 2 Digital Investigation 120.

2 Brian Carrier, 'Defining digital forensic examination and analysis tools using abstraction layers' and James R Lyle, 'NIST CFTT: Testing disk imaging tools' (2003) 1 Intl J of Digital Evidence; A D Irons, P Stephens and R I Ferguson, 'Digital Investigation as a distinct discipline: A pedagogic perspective' (2009) 6 Digital Investigation 82; Bradley Schatz, 'Digital Evidence: Representation and Assurance' (PhD thesis, Information Security Institute, Faculty of Information Technology, Queensland University of Technology 2007) <http://eprints.qut.edu.au/16507/1/Bradley_Schatz_Thesis.pdf>.

3 For instance, see *SWGDE Establishing Confidence in Digital Forensic Results by Error Mitigation Analysis* (1.5, 5 February 2015).

4 Erin Kenneally, *Gatekeeping Out Of The Box: Open Source Software As A Mechanism To Assess Reliability For Digital Evidence*, <www.vjolt.net/vol6/issue3/v6i3-a13-Kenneally.html>; Eric Van Buskirk and Vincent T Liu, 'Digital Evidence: Challenging the Presumption of Reliability' (2006) 1 Journal of Digital Forensic Practice 19; Lei Pan and Lynn M Batten, 'Robust performance testing for digital forensic tools' (2009) 6 Digital Investigation 71; SWGDE, *Recommended Guidelines for Validation Testing*, Version 1.1 (January 2009); Fred Cohen, Julie Lowrie and Charles Preston, 'The State of the Science of Digital Evidence Examination' in Gilbert Peterson and Sujeet Shenoi, (eds), *Advances in Digital Forensics VII*, 7th IFIP WG 11.9 International Conference on Digital Forensics, Orlando, FL, USA, 31 January – 2 February 2011 (Springer 2011); *Computer Forensic Tool Testing Handbook* (National Institute of Standards and Technology 2012); Jeremy Leighton John, *Digital Forensics and Preservation* (Digital Preservation Coalition 2012).

9.59 Automated tools are necessary to perform a forensic examination of a computer economically. However, the digital evidence professional should understand the process used by the tool to perform the relevant tasks. This is because it may be necessary to explain the process to a court, or the specialist may be required to carry out the analysis without the aid of a tool, because the use of a tool in any given situation may not be appropriate. These are issues that lawyers may well need to take cognizance of in the future.[1] For instance, it is not clear that practitioners themselves are familiar with some tools, and may question the worth of early versions.[2] This is because, it seems, that such tools are tested informally, rather than formally proven correct, and it has been suggested that such tools should be tested formally.[3] In an effort to enhance the veracity of evidence adduced from a forensic examination, it is becoming common practice within forensic laboratories to use what is known as 'dual tool' verification techniques. Simply put, an analyst will perform an examination using one piece of forensic software and, where data of potential relevance is identified, will use a second tool, produced by a different manufacturer, to perform the same examination and compare the results. If they match, more weight can be given to the accuracy of the data. However, it must be emphasized that such techniques are not a replacement for critical thinking or experimentation.[4]

1 For an example of where tools were the topic of judicial scrutiny in Australia, see *Bevan v The State of Western Australia* [2010] WASCA 101 and *Bevan v The State of Western Australia* [2012] WASCA 153. These cases are discussed in more detail in the chapter dealing with the presumption that computers

are 'reliable'. Interestingly, the lawyers nor the judges seem to have understood that the evidence remained stored on the SIM, even though the process of extracting the evidence was the subject of analysis.

2 Eoghan Casey, 'Network traffic as a source of evidence: Tool strengths, weaknesses, and future needs' (2004) 1 Digital Investigation 28.

3 James R Lyle, 'NIST CFTT: Testing disk imaging tools' (2003) 1 Intl J of Digital Evidence; Matthew Gerber and John Leeson, 'Formalization of computer input and output: The Hadley model' (2004) 1 Digital Investigation 214; Ibrahim M Baggili and Richard Mislan, 'Mobile phone forensics tool testing: A database driven approach' (2007) 6 Intl J of Digital Evidence; David Byers and Nahid Shahmehri, 'A systematic evaluation of disk imaging in EnCase 6.8 and Li En 6.1' (2009) 6 Digital Investigation 61; *SWGDE Recommended Guidelines for Validation Testing*, Version: 2.0 (5 September 2014).

4 See also Eoghan Casey, 'The increasing need for automation and validation in digital forensics' (2011) 7 Digital Investigation 103; Joshua I James and Pavel Gladyshev, *Challenges with Automation in Digital Forensic Investigations* (Digital Forensic Investigation Research Group, University College Dublin 2013), available at <http://arxiv.org/pdf/1303.4498.pdf>; mistakes were made in the case of Casey Marie Anthony in 2011, and one tool that was used did not give correct results, although once the designer was aware of the error, he informed the police immediately, Craig Wilson, 'Digital Evidence Discrepancies – Casey Anthony Trial' (11 July 2011) <www.digital-detective.net/digital-evidence-discrepancies-casey-anthony-trial/>; Tony Pipitone, 'Cops, prosecutors botched Casey Anthony evidence' (Clickorlando.com, 28 November 2012) <www.clickorlando.com/news/cops-prosecutors-botched-casey-anthony-evidence>; Baez and Golenbock, *Presumed Guilty* 46, 180–3, 211, 346–8, 365, 368–71, 400, 426–8; Jess Ashton and Lisa Pulitzer, *Imperfect Justice: Prosecuting Casey Anthony* (William Morrow 2011) 105, 239, 277, 291–2, 298, 315.

Copying the hard drive

9.60 Before obtaining access to a computer, it is essential that the investigator be familiar with the underlying operating systems, files systems and applications. By understanding the file systems, the digital evidence professional will be aware of how information is arranged, which in turn enables her to determine where information can be hidden, and how such information can be recovered and analysed. In order to establish answers to questions such as: 'who might have had access to a computer or system', 'which fils they would have been able to look at', and 'whether it was possible for an unauthorized outsider to obtain access to the computer from the Internet', the digital evidence professional should understand the nature of user accounts and profiles, and the control mechanism that determines which files a user is permitted to access once he is logged on to a system.

9.61 To acquire the data on a hard disk installed in a computer, an investigator will, in most cases, prefer to remove the hard disk from the computer and attach it to a specialist 'write protected' interface that is attached, in turn, to an 'imaging' device capable of copying the forensic image stored on the media on to a previously cleaned (and verified clean) storage device. Such interfaces are commonly referred to as 'write blockers', and the imaging capability may be performed by specifically designed imaging hardware or by a standard computer running imaging software. However, in some circumstances, removal of the hard disk from a computer may not be possible or advisable, in which case it is common to leave the hard disk installed in the host computer and obtain access to it using the procedures described in the following paragraphs.

9.62 To avoid altering any evidence on a computer, it is necessary to bypass the operating system. When the power supply is switched on, the basic input and output system (BIOS) will carry out a power-on self-test (POST) before looking for the

operating system. After the BIOS is activated and before the POST test has completed its cycle, it is possible to interrupt the process. Most computers are programmed to expect the operating system to be found on a floppy disk, hard disk, compact disc or a device attached to the Universal Serial Bus (USB). As a result, the system looks at these locations in the order set out in something called the Complementary Metal Oxide Silicon (CMOS) configuration tool. The CMOS chip retains the date, time, hard drive parameters, and other details relating to configuration while the main power is switched off. By looking at the CMOS tool between the POST test and the computer being fully powered up, the digital evidence professional is able to determine where the computer will look for the operating system: for instance, a floppy disk, a hard disk or a compact disc. By knowing where the computer is going to look for the operating system, the investigator is able to pre-empt the operating system on the computer and provide an alternate operating system from another disk. It is common for this alternative operating system to be a variant of the Linux operating system that is designed to allow storage devices to be viewed in 'Read Only' mode. By interrupting the normal boot up process in this way, the evidence on the hard drive remains intact and unaltered, thereby permitting the content to be copied in the state it was in when the computer was switched off. Various techniques and tools (such as an evidence acquisition boot disk) can be used to intercept this process, as the precise technique depends on the circumstances of each case.

9.63 Once the computer is booted from a suitable tool, the program can then do a sector-by-sector copy of the electronic evidence. Some tools will acquire the data and undertake an integrity check at regular intervals. There is a discussion in the electronic evidence field about whether some of the tools that undertake these tasks take an exact copy of the disk, even though all of the information is copied from the disk. One of the reasons is that data may be arranged in a different manner in a proprietary file format. Professor Casey suggests this is not as important as ensuring the integrity of the evidence is maintained, which must be correct. In addition, he also suggests that at least two copies be made with different tools.[1] From a practical point of view this may not always be possible, however, because of time constraints and the absence of storage media.

1 Casey, *Digital Evidence and Computer Crime*, 480.

9.64 A number of the forensic imaging tools have introduced the concept of 'Logical Evidence Files' which, instead of being an image of an entire hard disk, are copies of specific data (that is, the contents of a specific directory or directories). This technique has significant advantages where it is impractical to image an entire drive due to the amount of data required to be copied or because of time constraints. It should be noted that file hashing and image hashing techniques are still used to ensure the integrity of the data that is collected.[1]

1 Michael Cohen, Simson Garfinkel and Bradley Schatz, 'Extending the advanced forensic format to accommodate multiple data sources, logical evidence, arbitrary information and forensic workflow' (2009) 6 Digital Investigation S57; Da-Yu Kao, Shiuh-Jeng Wang and Frank Fu-Yuan Huang, 'SoTE: Strategy of Triple-E on solving Trojan defense in Cyber-crime cases' (2010) 26 Computer Law & Security Review 52.

Viewing the data

9.65 When the electronic evidence has been copied, the data can be viewed in raw format (examining the contents of the file in binary, hexadecimal or another format that displays the literal file contents as expressed in bits) or logically (using a viewer or program suitable for processing the file at hand). It is usually necessary to view the data through a tool. Human beings need the binary code, which resides on a disk or in a disk image, to be interpreted before the data can be viewed and interrogated in a sensible manner. In many tools for viewing raw data, the data can be viewed in hexadecimal form on one side of the screen and in plain text (ASCII or UNICODE) on the other side of the screen. Depending on the tool used, the data can be examined and analysed. For instance, a tool can recover slack or unallocated space[1] and compare files to determine if there are any differences to be observed.[2] Viewing data in logical view enables the user to examine it as represented by the file system. This way of looking at the data permits the user to analyse it in a different way, but it does not show the underlying information that is visible when using the physical method. Both forms of viewing data have their limitations, and it is also important to be aware that data can be misinterpreted. There is some debate about the best way of examining digital evidence, but the emphasis should be on verifying the accuracy of the evidence by using different tools.

1 Slack space is a part of a block or cluster of a filesystem that is used for another file, but that is not entirely overwritten by it. The block may then contain remnants of the file that was previously there. Unallocated space consists of blocks or clusters of the filesystem that were once used for a file, but, upon deletion of that file, are no longer referenced in the filesystems allocation table. They will contain the original content of the file until they are (fully) overwritten.

2 Note also that the volume of images that need to be reviewed and searched are increasing, and tools are being developed for this purpose: Paul Sanderson, 'Mass image classification' (2006) 3 Digital Investigation 190.

Recovering data

9.66 Increasing numbers of people delete the contents of their hard drives in computers in anticipation of legal action or after legal action has begun.[1] For instance, in the case of *L C Services Limited v Andrew Brown*,[2] Andrew Brown, the sales director of LC Services, was found to have broken the fiduciary duty he owed to LC Services. He also breached the terms of his services agreement and misused confidential information belonging to LC Services. It appeared that Mr Brown altered or re-installed the operating system on his computer on 1 October 2003, at the time the claimants were pursuing disclosure documents from the defendants. A digital evidence professional was subsequently able to retrieve the residue of the text of the relevant database in dispute, and the remains of a number of emails sent by Mr Brown. The content of these emails went to show that he was in breach of his fiduciary duties to LC Services.[3]

1 Ewa Huebner, Derek Bren and Cheong Kai Wee, 'Data hiding in the NTFS file system' (2006) 3 Digital Investigation 211; Dan H Willoughby, Jr, Rose Hunter Jones and Gregory R Antine, 'Sanctions for e-discovery violations: By the numbers' (2010) 60 Duke Law Journal 789.

2 [2003] EWHC 3024 (QB).

3 Bruce J Nikkel, 'Forensic acquisition and analysis of magnetic tapes' (2005) 2 Digital Investigation 8; Mayank R Gupta, Michael D Hoeschele and Marcus K Rogers, 'Hidden disk Areas: HPA and DCO' (2006) 5 Intl J of Digital Evidence.

9.67 There are several techniques that can be used to recover data that has been deleted. This can be done manually or through the use of tools, depending on the complexity of the problem faced by the specialist. For instance, some tools use a bit-for-bit copy of a disk to reconstruct the file system, including any files marked as deleted in the file allocation table, master file table or their equivalents. However, where files are fragmented and have been partially overwritten, it may be necessary to recover them by hand. A typical technique to recover deleted files (often called 'carving') involves searching unallocated space and swap files for such information as headers and footers. Although there are many types of file that can be recovered (carved) in this way with an appropriate tool, such as graphic files, word processing and executable files, recovery is limited to those files whose headers have not been deleted.[1]

1 Paul Alvarez, 'Using Extended File Information (EXIF) file headers in digital evidence examination' (2004) 2 Intl J of Digital Evidence.

Passwords and encryption

9.68 A number of tools are available that are capable of removing passwords and bypassing or recovering them. Some tools are available for guessing passwords if the encryption keys are small enough, and where it is not possible to defeat a password, it is sometimes possible to search for unencrypted versions of the data in other areas of the hard disk.[1] Passwords can be used simply to provide access control to unencrypted data, can be the 'key' that decrypts encrypted data, and can even be the 'key' that decrypts the actual key that is used to decrypt encrypted data. The methods used to bypass passwords or 'crack' the code needed to decrypt encrypted data are many and varied, but generally, stronger encryption algorithms and larger 'keys' mean that very long processing times are required to gain access to the data, if indeed they can be accessed. Depending on the processing power available, it may be impossible to reveal the passphrase or gain access to encrypted materials in a realistic time frame. The techniques used to attempt to obtain access encrypted or password protected data are discussed in the chapter on encrypted data.

1 Eoghan Casey, 'Practical approaches to recovering encrypted digital evidence' (2002) 1 Intl J of Digital Evidence; Christopher Hargreaves and Howard Chivers, 'Recovery of encryption keys from memory using a linear scan', *Proceedings of the 2008 Third International Conference on Availability, Reliability and Security* (2008) 1369–1376; Eoghan Casey, Geoff Fellows, Matthew Geiger and Gerasimos Stellatos, 'The growing impact of full disk encryption on digital forensics' (2011) 8 Digital Investigation 129.

Traces of evidence

Network connections

9.69 One of the most significant difficulties faced by digital evidence professionals with computers that are connected to a network such as the Internet, or a series of computers that are connected in an organization, is the possibility that a hacker or malicious employee might enter the system without authority and undertake a series of actions that causes an innocent person to be accused of doing something he did not do.[1] This is where data logs can help. Two types of logs, the application and system event logs, contain information about how users have used the computer. Scrutinising these logs, either manually or with a tool, can help to obtain a clearer picture about the activities that took place on the system, although consideration must be given to the

integrity of the logs themselves.

1 Srinivas Mukkamala and Andrew H Sung, 'Identifying significant features for network forensic analysis using artificial intelligence techniques' (2003) 1 Intl J of Digital Evidence; Bruce J Nikkel, 'Domain name forensics: A systematic approach to investigating an internet presence' (2004) 1 Digital Investigation 247; Bruce J Nikkel, 'Improving evidence acquisition from live network sources' (2006) 3 Digital Investigation 89; Eoghan Casey and Aaron Stanley, 'Tool review – remote forensic preservation and examination tools' (2004) 1 Digital Investigation 284; Omer Demir, Ping Ji and Jinwoo Kim, 'Packet marking and auditing for network forensics' (2007) 6 Intl J of Digital Evidence.

9.70 Note that logs may also be present at other levels in the network, such as on the fileserver, the Internet proxy server or the firewall. The availability of such logs may, however, vary a great deal. A typical problem in this area is the shared use of a single public IP address for Internet traffic by many different organizational users. These users will typically have their own, locally distributed, IP address. A setup like this is known as NAT (Network Address Translation) since it requires translation of the local user's (private) IP addresses to the public IP address and vice versa. Network based logs only rarely contain enough data to identify the individual user, however.[1]

1 Hein Dries-Ziekenheiner and Iljitsch van Beijnum, 'Allocation and Use of IP Addresses Study for the European Commission' (SMART 2010/14, December 2010) <http://bookshop.europa.eu/en/allocation-and-use-of-ip-addresses-pbKK0113063/>.

Logs, files and printing

9.71 In addition, when a user uses his computer, he leaves traces of his actions across a range of data logs and files.[1] A data log is capable of containing any type of data, depending on what the system is programmed to capture.[2] For instance, if a file is downloaded from the Internet, a date and time stamp will be added to the file to demonstrate when the file was downloaded on to the computer. When the file is moved, opened or modified, the time and date stamps will be altered to reflect these changes. In addition, the metadata can also help provide more information about the file, such as the location it was stored on the disk, the printer on which the file was printed and the time and date the file was created. When a file is printed, the computer tends to store the print job in a temporary file, before it is sent to the printer when the printer has the capacity to print the file. Once the command to print has been passed to the temporary store, the user can continue to work with the application – for instance, they can continue to type a new document while the previous document is waiting to be printed. The temporary print store retains valuable information, such as the name of the file to be printed, the type of application used, the name of the printer, the purported name of the person whose file is to be printed, and the data itself. In addition, there is a date and time stamp added to the file to show when the file was printed. It should be noted, however, that the date and time stamp can be altered, which means it is important to ensure that the time and date stamp is corroborated by other methods.[3]

1 In relation to intrusion detection systems, see Peter Sommer, 'Intrusion detection systems as evidence' [2002] CTLR 3, 67; Vlasti Broucek and Paul Turner, 'Intrusion detection: Issues and challenges in evidence acquisition' (2004) 18 International Review of Law, Computers & Technology 149; Jean-Marc Dinant, 'The long way from electronic traces to electronic evidence' (2004) 18 International Review of Law, Computers & Technology 173.

2 Erin E Kenneally, 'Digital logs – proof matters' (2004) 1 Digital Investigation 94.

3 Karen Kent and Murugiah Souppaya, *Guide to Computer Security Log Management* (Special Publication 800-92 2006) at <http://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication 800-92.pdf>, 2.1.3 fourth bullet point.

Use of the Internet

9.72 When a person obtains access to the Internet, a range of data is created and retained on a computer, including the websites that have been visited, the contents a user has viewed and what data sources he has obtained access to.[1] Some systems, both in the network and on customer premises, also include a log of the times and dates of the Internet session and details of the device or connection that was used (such as the modem, network card or physical network port in the access network). With more services available online, it is important to be able to rely on information provided to the investigation by network operators. Typical information requests involve IP addresses, subscriber details and possibly payment information. Internet access logs may, furthermore, provide information as to where and how users were connected to a service, and may identify others involved in the same investigation.

1 Yeong Zee Kin, 'Computer misuse, forensics and evidence on the Internet' (2000) 5 Communications Law 153; Vivienne Mee, Theodore Tryfonas and Iain Sutherland, 'The Windows Registry as a forensic artefact: Illustrating evidence collection for Internet usage' (2006) 3 Digital Investigation 166.

9.73 Finally, it is interesting to observe that CCTV systems are gradually being replaced by systems that use Internet Protocol technologies (IP) and wireless IP, which will in turn cause additional expense and increase the legal complexity (where the camera is taking images in one country, and they are being recorded or stored in another country) in obtaining access to such systems for the purposes of litigation or criminal proceedings.[1] The types of information available include those noted below.

1 Fanny Coudert, 'Towards a new generation of CCTV networks: Erosion of data protection safeguards?' (2009) 25 Computer Law & Security Review 145.

9.74 Browser cache When viewing a page on the Internet, the browser retains and takes copies of all the elements that make up the page, such as graphics and HTML text, called a cache. The computer gives the page a date and time stamp at the time the page was downloaded. The reason for doing this is that when the page is visited again, the cached file is used by the computer in place of obtaining access to the same page online and the date and time stamp is subsequently updated. Another item of information created and logged in some browser history databases is the number of times a web page was visited. It must not be assumed, however, that just because the computer has recorded certain types of web pages that the user actually viewed such pages. This is because some websites, in particular those promoting pornography, will redirect a browser to different websites, and may even make changes to the computer that the user has not authorized.[1] It is possible to recover these cached files, even if they are deleted. Recovered files can provide such information as when the computer was used to obtain access to web based email, when sites were visited and if purchases were made or financial transactions undertaken.

1 Daniel Bilar, 'Known knowns, known unknowns and unknown unknowns: Anti-virus issues, malicious software and internet attacks for non-technical audiences' (2009) 6 Digital Evidence and Electronic Signature Law Review 123.

9.75 Cookies Many websites keep a track of visits by users to their sites by placing this information in files on the users' computers called cookies. If cookies have not been disabled, the information in the cookie directory can help with an investigation. As for websites included in the temporary cache file, it does not follow that just because there is a cookie on the computer that a user necessarily went to all of the websites included in the cookie directory. Some advertisements on a website may place a cookie on the user's computer, even though the user did not click on and view the particular website. Further, where the user's browser has been redirected without his permission, cookies can be added to the directory without the knowledge of the user.

9.76 Private browsing, VPN proxies and TOR In order to provide Internet users with more privacy, several browser manufacturers have introduced 'incognito modes' or 'private browsing' modes in their browser software. In this mode, no Internet history, cache entries or cookies (or any other artefacts) remain after the Internet session. This means it will be harder (if not impossible) to retrieve a reliable indication of a user's Internet usage and surfing behaviour. In practice, other systems such as access logs at service providers or websites may still be able to identify the user by his IP address.

9.77 In order to further enhance user privacy and anonymity, services such as TOR (The Onion Router) and VPNs (Virtual Private Networks) are available that allow users to hide the origins of their connection to the services they use. In the case of The Onion Router this is achieved through a network of nodes operated by volunteers who anonymize connections to the Internet by providing a route across three or more anonymous nodes (including an entry and exit node, as they are called) on behalf of a TOR user. Since no logging is kept at any of the intermediary TOR nodes, this assures a relatively high level of anonymity. Similarly, VPNs and proxies can be used to connect to the Internet via a predetermined 'hop' in the network. Provided the VPN origin is not logged, this may effectively make tracing users by their network addresses impossible. Note that the use of other information and identifiers is still possible so that various other measures may still reveal the users' actual names, addresses and Internet activities.

9.78 Email and instant messaging Email has become a dominant method of communication for the vast majority of organizations, although in its place, the 'chat' feature is probably used widely by individuals on smartphones. As a result, a great deal of evidence can be discovered from email correspondence. Some software programs store email in plain text files, while others use proprietary formats that will require the digital evidence professional to use a number of tools in order to read the messages. Other email systems utilize online storage only and leave very little communications data on the computer hard disk. It is sometimes possible to recover email messages that have been deleted but have not been removed from the email files. Where it is impossible or difficult to restore emails from a single computer, it might be possible to track email traffic through the network it has travelled.[1] Organizations are beginning to recognize the importance of their email correspondence, and many larger organizations have archives of email communications that can be investigated in the event of e-disclosure or e-discovery requests.

1 Eoghan Casey, Troy Larson and H Morrow Long, 'Network analysis' in Casey (ed), *Handbook of Computer Crime Investigation Forensic tools and technology* 234–9.

9.79 Instant messaging, in the meantime, has become the default method of communication for many people. This presents problems for the investigator. It is not only used on local desktop systems (where this technology is increasingly also used in business environments) but it has also seen a major surge in use on mobile devices in recent years. Due to the Snowden revelations, many instant messaging programs currently in widespread use have introduced end-to-end encryption, meaning that intermediaries do not have access to plain text messages, but merely to an encrypted version. Each connected device has a unique public key and a private key that is unknown to the intermediary. In practice this means that the only place where such communications can be accessed and decrypted to a readable format is on the telephone or the end user's device.

9.80 Mobile applications that are used for instant messaging typically include the ability to send photographs and videos. Social networks and mobile Internet messaging have become the default communication method used by children.[1] This creates an increased workload for investigators of child-abuse related cases, especially where they may need to view a home computer, as well as a multitude of other devices, to help determine why a child might have left home, or how he got into contact with a certain adult, for instance.[2] Another challenge is that these programs increasingly offer features that allow the user to set a destruction timer on any images he sends. Images, therefore, are no longer stored on the filesystem of the device or telephone by default, but only temporary copies are displayed for a short period of time, after which they are deleted. This leaves fewer artefacts and creates further challenges in criminal investigations involving abusive images and children, particularly in relation to practices such as sexting, the sending of sexual images and messages, and grooming, where adults lure children typically for sexual abuse by acting as persons of the same age.

1 Sonia Livingstone, Leslie Haddon, Anke Görzig and Kjartan Ólafsson, *Risks and safety on the internet: the perspective of European children: full findings and policy implications from the EU Kids Online survey of 9-16 year olds and their parents in 25 countries* (LSE 2011) <www.lse.ac.uk/media%40lse/research/EUKidsOnline/EU%20Kids%20II%20(2009-11)/EUKidsOnlineIIReports/D4FullFindings.pdf>.

2 Harlan Carvey, 'Instant messaging investigations on a live Windows XP system' (2004) 1 Digital Investigation 256; Mike Dickson, 'An examination into MSN Messenger 7.5 contact identification' (2006) 3 Digital Investigation 79; Mike Dickson, 'An examination into Yahoo Messenger 7.0 contact identification' (2006) 3 Digital Investigation 159; Paul Sanderson, 'Identifying an existing file KaZaA artefacts' (2006) 3 Digital Investigation 174; Mike Dickson, 'An examination into AOL Instant Messenger 5.5 contact identification' (2006) 3 Digital Investigation 227; Jessica Reust, 'Case study: AOL instant messenger trace evidence' (2006) 3 Digital Investigation 238.

9.81 Voice over Internet Protocol (known as VoIP) is another computer-to-computer technology that has expanded rapidly, and will need to be considered when conducting an investigation.[1] Contrary to the old telephony system (often referred to as POTS or Plain Old Telephone System), Internet based calls can be made fairly anonymously and telephone numbers can easily be spoofed, especially those of the party initiating the call.[2] This makes telephone numbers increasingly unreliable as identifiers. The risk of wrongfully attributing the source of a call on the basis of its originating telephone number has increased manifold, especially since many VoIP providers allow spoofing of outbound calls as a service feature, and special services have emerged that specialize in spoofing calls for various purposes. In both cases the connection will be encrypted which means that the data packets flowing between the caller and the recipient of a

VoIP call is not in decipherable voice form and cannot be reconstructed to meaningful evidence, if intercepted.

1 Xinyuan Wang, Shiping Chen and Sushil Jajodia, 'Tracking anonymous peer-to-peer VoIP calls on the Internet', Proceedings of the 12th ACM conference on Computer and Communications Security (2005), 81–91.
2 Richard Clayton, 'Can CLI be trusted?' (2007) 12 Information Security Technical Report 74, <www.cl.cam.ac.uk/~rnc1/cli.pdf>.

Reporting

9.82 The findings, and any conclusions made by the digital evidence professional, will be set out in a report. Whether prepared for criminal or civil proceedings, the report should include a range of information that is pertinent to the case, including, but not limited to:

(i) Notes prepared during the examination phase of the investigation.

(ii) Details about the way in which the investigation was conducted.

(iii) Details about the continuity of custody.

(iv) The validity of the procedures used.

(v) Details of what was discovered, including, but not limited to:

(a) Any specific files or data that were directly related to the investigation.

(b) Any further files or data that may support the conclusions reached by the specialist. This will include the recovery of any deleted files and the analysis of any graphic files.

(c) The types of search conducted, such as key word searches, and the programs searched.

(d) Any relevant evidence from the Internet, such as emails and the analysis of websites visited and log files.

(e) Indications of names that might demonstrate evidence of ownership of software, such as with whom the software was registered.

(f) Whether there was any attempt to hide data in any way, and if so, what methods were used.

9.83 Professor Casey refers to the following principles: observation, hypothesis, prediction, experimentation/testing and conclusion.[1] Following from these principles, the report needs to reflect how the examination was conducted and what data were recovered. It may be that the digital evidence professional will have to give evidence about her conduct of the examination and the validity of the procedures and tools used. Essential to any report will be the conclusions reached by the professional. Where an opinion is offered, the opinion should set out the basis of the evidence. Consideration should also be given to rates of error, including the origin and timing of events that had been recorded, whether the digital evidence professional took care when reaching conclusions where data were lost, whether the professional was aware that digital evidence can be fabricated, and whether the professional evaluated the evidence based 'on the reliability of the system and processes that generated the records'.[2]

1 Casey, *Digital Evidence and Computer Crime*, 204.
2 Eoghan Casey, 'Error, uncertainty, and loss in digital evidence' (2002) 1 Intl J of Digital Evidence.

9.84 As pointed out by Professor Sommer, it is important to be aware that digital evidence professionals have to use a variety of techniques to cope with the wide diversity of hardware and software encountered. Reliability is one factor to take into account. Another factor is the degree of reliance on the conclusions reached by a digital evidence professional. The digital evidence must be interpreted, and care should be taken to ensure the underlying rationale is sustainable.[1] Assumptions should not form part of any report by a digital evidence professional, as occurred in some cases relating to the investigations by the UK police under the name Operation Ore. In this case, police forces in the UK investigated and prosecuted over 7,000 people for offences relating to abusive images of children, and secured over 2,000 convictions. This operation was instigated after the conviction of Thomas and Janice Reedy (the Landslide trial, named after their company) in the United States for operating a web site selling access to abusive images of children.[2] After the trial, a copy of the database recording details of the payments received by Landslide was shared with a number of police forces across the world. This information formed the initial evidence for the purposes of the investigations that subsequently took place.

1 Peter Sommer, 'Digital footprints: Assessing computer evidence' [1998] Crim LR Special Edition, 65 and 69.
2 *United States of America v Reedy*, 304 F.3d 358 (5th Cir. 2002).

9.85 There was evidence to suggest that stolen credit card numbers were used to steal money by 'buying' access to the illegal websites hosted by Reedy, who tried to prevent this without success.[1] Some of those prosecuted claimed that they did not use their credit cards to obtain access to abusive images of children, as in the case of Dr Paul Grout. No abusive images of children were found on his computers. He produced alibi evidence to demonstrate that at the time of the alleged links to the Landslide website, he was not at a computer terminal. The case was withdrawn from the jury.[2]

1 Duncan Campbell, 'Sex, lies and the missing videotape', *PC Pro* (June 2007), 18–21; Supplementary memorandum by Mr Jim Gamble dated 1 June 2007 submitted to the Science and Technology Committee – Fifth Report (Session 2006-07, 24 July 2007) (the evidence is published in Vol II (HL Paper 165-II)), where Mr Gamble challenges some of the assertions made by Mr Campbell.
2 'Invisible predator' *BBC, Inside Out – Yorkshire & Lincolnshire*, 4 October 2004.

9.86 On occasions, it was also assumed that if a credit card number was in the Landslide database, the person whose number it was had therefore paid for abusive images of children. Brian Cooper used his credit card to buy bicycle parts from a US website. His card details were obtained by Akip Anshori, an Indonesian, who successfully subscribed to the Landslide website until Mr Cooper alerted his credit card provider to the unauthorized payments. The police failed to find any abusive images of children on his computers.[1]

1 Campbell, 'Sex, lies and the missing videotape', 19.

9.87 Jeremy Clifford was charged with making and being in possession of indecent images of children. The images were found in the temporary cache folder with random names such as 'FX7RA'. Such images generally appear as advertisements, and the user will not necessarily have clicked on them, nor will he be aware that they are on the machine. At his trial, Mr Clifford was acquitted when the prosecution offered no evidence. He failed in his first legal action for malicious prosecution and misfeasance

in public office,[1] although his appeal succeeded,[2] and the police were subsequently found liable.[3] The police and the digital evidence professional had made a number of assumptions about the evidence they found, and Mr Justice Mackay had cause to address the nature of the technical evidence, which bears repeating in full:

> The Expert Issue
>
> 67. As a postscript to the above I should give my reasons for the finding that I do not accept the argument that the claimant was an habitual seeker out of IIOC [*indecent images of children]* from 1999–2005. This was based on expert evidence and their evidence was limited to this issue.
>
> 68. As to the Tiny computer itself there was no evidence that it had ever been used for such a purpose. The agreed evidence of the computer experts (Mr Fellows for the defendant and Mr Campbell for the claimant) included these statements
>
> > 'We found no evidence on the computer which indicated that the user searched the web for IIOC on any occasion… the computer was not "cleaned" or wiped so as to remove deleted records or information…the presence on the computer of a significant number of current and deleted sexually explicit adult images is consistent with the user or users of the computer having browsed adult websites prior to 11 February 2001. If the user or users had at these times browsed sites offering or supplying IIOC, or sites of ambiguous or similar character and which would be likely to launch pop ups containing IIOC, it would be very likely that some of these images would have survived and have been recovered'.
>
> 69. Mr Fellows' opinion supporting the defence proposition was not based on data found in this or any other of the claimant's computers but primarily on his general experience. He said that broadly speaking users of the Internet remained untroubled by advertising material of this nature if they are not interested in such material, and even if the user visited adult sites generally speaking they would remain untroubled by such pop ups. There was ample evidence of adult pornography being sought out by this computer but there were no IIOC images.
>
> 70. Mr Campbell says that adult site users, as this claimant admittedly was, might indeed suffer or attract pop ups of this nature. He had very considerable experience of such cases and Landslide cases in particular, albeit acting exclusively for the defence. His opinion is that the IIOCs found on the computer were the result the malware launched from adult pornography website pages. He has possession of the entire Landslide database which he has perused and finds evidence of massive and widespread credit card fraud on it, which indeed marked the beginning of the end for Landslide causing it to be closed down in 1999, and it was his opinion that such fraudulent use cards and/or personal information was potentially responsible for the Landslide entries in 1999 incriminating the claimant.
>
> 71. In cross examination Mr Fellows modified his position to this extent. It had never been his evidence that the claimant habitually trawled the net for IIOC. There is evidence, he thought, that the claimant used to visit child pornographic sites 'albeit not deliberately'. There are indications of visits to 'Lolita' and 'Pre-Teen' which 'may or may not have been deliberate' and his accessing of Lolita.2000 may not have been deliberate.
>
> 72. I prefer Mr Campbell's position on this issue particularly in the light of Mr Fellows concessions recorded above.
>
> 73. Secondly Mr Fellows pointed to an email from jezz1@aol.com (the claimants email account) to world-bill.com (a payment processing site) concerning a site 'Lolita.2000' (which no longer exists and the content of which nobody knows)

> and the password needed to access it. Mr Campbell showed me the full picture here. This email was part of a series of ten which began with a bizarre email from world-bill to the claimant on 7 December telling him and 23 other alleged 'subscribers' that their 'subscription' had been renewed. It did not say to what sites or for what charge. The claimant replied that he did not understand as he had not requested any account to be renewed. The third, fourth and fifth emails are all from world-bill.com and are incomprehensible. They make no reference to any particular website. The sixth told him that his subscription to Lolita.2000 was now active (this is the first time the site has been named) and it gave details including his username and password. The seventh is the only other email the claimant sent in this series and it said 'password is not working please advise'. The eighth is some form of automatic response from world-bill to the claimant. The ninth sent him a different username and password and the tenth cancelled his subscription.
>
> 74. Apart from his response, perhaps out of curiosity, indicating the password was not working, the claimant seems to have been the passive recipient of this strange series of emails. More to the point there is no evidence what Lolita 2000 contained as a site. I am unable to find that the claimant did anything to provoke this correspondence. It does nothing to support the proposition advanced by the defendant.
>
> 75. As to the Landslide transactions of 1999, only one is now capable of being shown to have been a subscription to a child pornography site namely 'Nympho'. The claimant's assertion in his 2003 interview that there had been fraudulent use of his card and details was one which curiously he himself did nothing to support by way of investigation or complaint, at least that he can now recall. It is however supported to a significant extent by material within the Landslide database which the expert Mr Campbell possesses and has searched. His searches show that three payments of the six were in fact refunded on 26 August 1999 which includes that in respect of the subscription to the site 'Nympho'. The claimant says that that evidence is also in the hands of the prosecution albeit in the huge file that constitutes the Landslide database but which is evidence capable of being searched electronically, as Mr Campbell has done.
>
> 76. Therefore there is no evidence to support the proposition advanced as a matter of fact by the defendant in this case.[4]

1 *Clifford v The Chief Constable of the Hertfordshire Constabulary* [2008] EWHC 3154 (QB).
2 *Clifford v The Chief Constable of the Hertfordshire Constabulary* [2009] EWCA Civ 1259.
3 *Clifford v The Chief Constable of the Hertfordshire Constabulary* [2011] EWHC 815 (QB).
4 [2009] EWCA Civ 1259, [67]–[76].

9.88 Great care must be given to the nature of the technical evidence, as demonstrated by the case of *O'Shea v The Queen*, a case that also centred on the Landslide database.[1] The case had been publicised by the media as a public enquiry into the entire operation conducted by the police. It was not. It was an appeal against conviction by one man on the main ground that new evidence from one Bates, described as a computer expert, based on a forensic examination of the Landslide records, suggested that a third party had misappropriated the appellant's identity. Bates had also suggested that evidence that the transactions recorded against O'Shea's name had been made from a computer with a Freeserve IP address (O'Shea had an account with Freeserve as his ISP) did not prove that O'Shea's computer had been used to obtain access to the Landslide computer, because an IP address can be disguised.

1 [2010] EWCA Crim 2879.

9.89 The Court of Appeal considered the submissions carefully. On the issue of the disguised IP address, Stanley Burnton LJ noted that this matter was canvassed at the trial. The digital evidence professional for the Crown had agreed that it was possible to disguise an IP address, but indicated that it was not possible to assume an IP address allocated to someone else. O'Shea's other expert witness had accepted the Crown's evidence in this regard. Stanley Burnton LJ also noted that Bates had been instructed before the trial, prepared a report for the trial, and was present at the trial. Bates' evidence was considered with caution, because he failed to raise this issue at the trial, denied having signed a confidentiality agreement relating to another case until it was produced to him in court, and in 2008 he was convicted at Leicester Crown Court for perjury for misrepresenting his qualifications when giving evidence. The members of the Court of Appeal held that there was no evidence to support Bates' suggestion that the Landslide webmaster had access to the appellant's personal data that were used in the transactions, that there was no evidence to prove that the hypothetical fraudulent webmaster had obtained access to the Freeserve proxy servers to assume the appellant's identity, and noted the appellant had checked his credit card statements regularly and not challenged these transactions (he had challenged the debiting of his credit card account in relation to other amounts that were similar to those in question in this case).[1] Describing this additional evidence as 'mere assertion, unsupported by any published or other material or any reasoning,'[2] the members of the Court of Appeal concluded that the appellant's conviction was safe and dismissed the appeal.

1 [2010] EWCA Crim 2879, [50]–[59].
2 [2010] EWCA Crim 2879, [43].

Analysis of a failure

9.90 A prosecution in Wales in 2015 offers a good case study to demonstrate what can go wrong when the police do not conduct a careful investigation, and the prosecution's failure to understand the weakness of the evidence upon which the charges are preferred. A number of nurses working at the Princess of Wales Hospital in Bridgend were indicted on charges relating to alleged falsification of patient notes regarding blood glucose levels. The evidence is discussed in detail by Professor Thimbleby, an expert witness for the defence,[1] where he outlined the correct procedure to be observed by the nurses:

> 1. Find a glucometer;
>
> 2. The nurse then identifies themselves [sic] to the device (by scanning the barcode on their [sic] staff card or by typing their [sic] ID);
>
> 3. The patient ID is scanned from a barcode or typed;
>
> 4. The patient is pricked and a drop of blood placed on a test strip;
>
> 5. The test strip is inserted in the glucometer;
>
> 6. The glucometer displays the blood glucose level (or possibly an error);
>
> 7. The nurse may then take immediate action to address any clinical issues;
>
> 8. The nurse then 'contemporaneously' writes down on the paper patient notes the time and reading.
>
> 9. One further step, that has no immediate clinical significance, is that the glucometer must be placed in a dock, and then its data will be automatically uploaded to [the] central database in the hospital.

1 Harold Thimbleby, 'Cybersecurity problems in a typical hospital (and probably in all of them)' Safety-Critical Systems Club (forthcoming 2017).

9.91 The central record system had no records of many of the tests the nurses had written on the paper notes for each patient. Because of this discrepancy, the police concluded that the nurses had written down fictitious readings, and had not bothered to do their job. As an aside, nurses would not necessarily undertake the actions as set out above, because of problems with the software. Sometimes it was difficult for the software to read the patient's identity number. It turned out that a practical solution was to type 000 on the glucometer keyboard, or for the nurse to scan her own barcode in order for the glucometer to accept the data to be input as a valid patient, or to manually type in the name of the patient – but this action would not prevent the nurse from misspelling the patient's name. The glucometer accepted both of these methods of getting around the failure of the software code, and would give a correct blood glucose reading. However, the hospital system rejected this data, the consequence of which required manual intervention for the data to be added to the central database – which might not happen, or might introduce further errors.

9.92 On analysing the prosecution evidence – which was a CD of Excel spreadsheets and, on a later date, data logged on blood glucometers in XML files – it was discovered that the relevant data was not present. The prosecution asserted that because data was not present, it followed that the nurses had fabricated doing actual tests, because if they had actually done the tests, the data would be present in the spreadsheets. Professor Thimbleby accepted that this was a logical possibility, but he thought it far more likely that there was a simpler explanation to the IT problems, especially because the system allowed administrators to make arbitrary changes to data. Upon analysing the evidence, a number of issues arose: more than 20 per cent of the data had an error flag; a comment field on each test said 'Wrong patient' for only 2 entries and nothing at all for over 130,000 entries; a 'reviewed' flag was false for almost all of the entries; staff names occurred with many implausibly close variant spellings, and many identical staff names occurred with multiple numeric identity numbers. This caused Professor Thimbleby to conclude that the problems were more complex than the prosecution portrayed. It was suggested that the database was not well managed, and therefore might not be reliable as evidence to show the nurses' complicity.

9.93 The actions of the police in copying the data were also suspect. The police maintained that they used forensic methods to make the copies of the database. There had been manual intervention: the hospital database was in SQL and it differed from some of the Excel spreadsheets, which strongly suggested that a manual process had been used to create or edit the spreadsheets. In addition, it was impossible to tell whether rows or columns had been deleted, had been edited, or had never existed in the Excel spreadsheets. The police copied the database at the hospital on to a USB stick, and only then digitally signed the data. This action was too late. The police should have made a signed copy of the original SQL database, not a copy of the Excel data created manually from the SQL database and copied to a USB stick. Furthermore, the police seized several blood glucometers from the ward and presented the data on them as evidence. However, they failed to seize at least two other glucometers that had been docked on the relevant ward over the period of the alleged fabrications, and the database only indicated where glucometers were docked, not where the blood glucose

tests were made. In addition, the devices moved around the hospital, and as there was no inventory, the police would not have known which device to seize.

9.94 Professor Thimbleby indicated that it was possible that the alleged failure to measure the blood glucose levels might still be stored on a glucometer somewhere in the hospital, and that the data has not been added to the central database because the device was still waiting to be docked. One XML file showed a four year gap between a measurement being taken and the data transferred to the database. The alleged incidents happened less than four years before the trial, so it was possible that the missing data had yet to reach the central database. There was also evidence that the Excel spreadsheets were of low quality; the software code on the glucometers was so poor that there was no reliable connection between the glucometers and the SQL database. Other points of failure included: the fact that a glucometer might lose data itself; that the device might not be docked; a device might be physically lost or returned for repair; a docking might fail, whether because of manual interference in the ward or for technical issues such as Internet connectivity problems, new servers and so on; the glucometers store about 2,000 readings, yet the database showed that they were used for nearly 5,000 tests; once docked, the data could take days to get through to the main database, and manual intervention was required for some data to be uploaded, but there was no evidence to demonstrate that manual intervention occurred.

9.95 The prosecution needed to prove that it was the failure of the nurses to input data that caused the data not to be present in the central database – that is, the absence of data proved fabrication, rather than any other possibility. The police and the prosecution lawyers assumed that the glucometers and hospital IT systems were reliable, even though they knew the systems required human intervention. The police did not question the management of the data, and there was no evidence about the day-to-day management of the data. The prosecution also claimed that the devices were accurate as blood glucose meters. This was not relevant. The relevant issue was whether the glucometers reliably transmitted test data to the patient record system. It did not appear that the police or the prosecution bothered to research this topic – if they had, they would have discovered a number of relevant articles that included reference to the issues noted by Professor Thimbleby regarding the practical problems of the device and getting the data to the central computer.[1] The judge concluded that the prosecution evidence was unreliable and was therefore excluded.[2] The prosecution response was to offer no evidence. There the matter ended for the nurses who entered pleas of not guilty.[3]

1 Ksenia Tonyushkina and James H Nichols, 'Glucose Meters: A Review of Technical Challenges to Obtaining Accurate Results' (2009) 3 Journal of Diabetes Science and Technology 971; Suzanne Austin Boren and William L Clarke, 'Analytical and Clinical Performance of Blood Glucose Monitors' (2010) 4 Journal of Diabetes Science and Technology 84; James H Nichols, 'Blood Glucose Testing in the Hospital: Error Sources and Risk Management' (2011) 5 Journal of Diabetes Science and Technology 173; David C Klonoff, 'Point-of-Care Blood Glucose Meter Accuracy in the Hospital Setting' (2014) 27 Diabetes Spectrum 174.

2 *R v Cahill and Pugh*, The Crown Court at Cardiff, ruling by HHJ Crowther QC, 14 October 2015 – Stephen Mason has been furnished with a copy of this ruling, but it is not available publicly.

3 'Nurses cleared of wilful neglect at Princess of Wales Hospital in Bridgend' *South Wales Evening Post* (Swansea, 14 October 2015) <www.southwales-eveningpost.co.uk/nurses-cleared-wilful-neglect-princess-wales/story-27983645-detail/story.html>; 'Princess of Wales Hospital nurse neglect trial collapses', *BBC News* (14 October 2015) <www.bbc.co.uk/news/uk-wales-south-east-wales-34527845>.

Anti-forensics and interpretation of evidence

9.96 As with all fields of forensic analysis, computer forensics is part of a continuous race of catch-up between investigators and criminals. Just as criminals quickly started to wear gloves once fingerprint evidence had reached the awareness of the wider public, computer criminals too began to use tools to hide or alter the traces of their activities. Anti-computer forensics has become the term for the possible countermeasures that criminals may take to prevent, delay or invalidate computer forensic efforts, a problem increasingly recognized by the research community.[1] Deletion of data as a classic anti-forensic technique may serve as an initial example to illustrate some of the issues computer crime investigations are increasingly confronted with. In the early days of the Internet, software that securely wiped data from all parts of the computer was the preserve of the experts, or governmental organizations with special security needs. Today, tools that irretrievably delete files are now easily obtainable for free from various sources, and can be used quickly and reliably even by comparatively computer-illiterate users.[2] This example not only illustrates the proliferation of anti-forensic tools, it also highlights some of the complexities that are involved. Most anti-forensic tools are 'dual nature' tools, just as many hacking tools are. They have legitimate uses and are often even officially recommended, if not legally mandated, for instance, to protect the security and privacy of sensitive data. Computer software is regularly 'purpose neutral'. In other words, what works as a protection against criminals trying to obtain access to credit card details also works as a protection from the police trying to obtain access to private emails; what works for system administrators seeking to detect misuse of a computer by an employee also works for criminals obtaining access to commercially sensitive secrets. This has implications for the legal responses to anti-computer forensics, and also for the probative weight of evidence respecting any counter measures that were used by a suspect, and is further discussed below.

1 Chris B Simmons, Danielle L Jones and Lakisha L Simmons, 'A framework and demo for preventing anti-computer forensics' (2011) 11 Issues in Information Systems 366; R Harris, 'Arriving at an anti-forensics consensus: Examining how to define and control the anti-forensics problem' (2006) 3 Digital Investigation S44.

2 Andy Jones and Christopher Meyler, 'What evidence is left after disk cleaners?' (2004) 1 Digital Investigation 183; Laurent Simon and Ross Anderson, *Security Analysis of Android Factory Resets*, <www.cl.cam.ac.uk/~rja14/Papers/fr_most15.pdf>.

9.97 As noted above, the social context is a crucial determinant for the interpretation of electronic evidence. In the early days of the Internet, finding that a suspect had acquired the specialist knowledge necessary to operate (or maybe even write) the software for a cleaning tool could be prima facie evidence that he had tried to hide traces of illegal activity. This inference is no longer sound, because secure cleaning of deleted data has become a standard operating procedure in many organizations to prevent data security breaches, and default settings on popular free tools such as CCleaner allow the effortless routine destruction of deleted files every time a computer is shut down.

9.98 The legal system and police investigators have reacted in several ways to this new reality. One approach is through technology – developing new investigative tools that either look for other types of data not yet protected by counter measures, or are in some other way capable of undoing the damage of anti-forensic tools. However, this

need to react rapidly to developments in the anti-computer forensic field can cause problems for the legal system, where rules on the admissibility of scientific evidence often require extensive testing and acceptance in the scientific community, supported by publication in peer-reviewed journals, together with robust methods of calibration, standardised procedures, accepted minimum criteria for training and proficiency with the new tools.[1]

1 For the US, see Christopher V Marsico, *Computer evidence v. Daubert: The coming conflict* (2004) CERIAS Tech Report 2005–17; the issue was also discussed in the context of anti-forensics and the use of the 'Evidence Eliminator' programme in *State of Ohio v Starner*, Slip Copy, 2009 WL 3532306 (Ohio App. 3 Dist.), 2009 -Ohio- 5770; Barbara Guttman, James R Lyle and Richard Ayers, 'Ten years of computer forensic tool testing' (2011) 8 Digital Evidence and Electronic Signature Law Review 139; Computer Forensics Tool Testing (CFTT) Project, <www.nist.gov/content/computer-forensics-tool-testing-cftt-project>; DigitalCorpora.org, <http://digitalcorpora.org>.

9.99 What is important to note for criminal prosecution purposes is that electronic evidence can serve a dual purpose: it can either directly support the prosecution's case, or it can be indirect evidence that the suspect took actions to hide some form of criminal activity – which in turn may also be direct evidence that he committed one of the various statutory offences that have been created to prevent the destruction or spoliation of data.

9.100 With all this in mind, the following is a short overview of the various approaches to anti-computer forensics, and the effects they have on the availability, reliability and interpretation of electronic evidence. Anti-computer forensics are understood here as any technique, hardware tool or software that prevents or delays the forensic analysis of a data carrier, and negatively affects the existence, amount, authenticity or quality of evidence from a computer. There are at least five different subcategories of anti-forensics: data destruction, data tampering, data hiding, trail obfuscation and attacks against the computer forensic tools themselves.

Data destruction

9.101 Data destruction is the most obvious and most widely discussed anti-forensics measure, and has created a considerable legal and technological debate.[1] Unlike a physical object or piece of paper that can be destroyed effectively, it is much more difficult to completely obliterate a document in electronic format. All a user does when he clicks the 'delete' icon on a computer is, in general terms, to remove the pointer to the data. The document or data remains, and it is possible to retrieve this data in certain circumstances, even if it is partly overwritten.[2] However, disk cleaning utilities that overwrite or 'shred' data have become increasingly available and easy to use for even unsophisticated users. These software-based tools write patterns of pseudo-random combinations of 1s and 0s (in other words, meaningless data) on to all of the sectors on a hard drive. This also includes a setting to wipe free space or unallocated or 'slack' space, which is where older 'deleted' data often reside. Slack space occurs when data is split between clusters on the hard disk. As files only rarely and by chance fill every cluster up, some space remains. (Think of a collection of standard-sized shoeboxes in which you store documents. If you 'delete' a file in such a shoebox, it simply allows you to put a different document in the box by placing the new document on top of the space previously occupied by the older document, but if the new document is smaller, parts

of the old file remain in the 'slack space' and can be recovered). Cleaning software also deletes much of the metadata that accumulates from using the computer – it wipes and cleans old file entries, recently used file lists and many other things including custom locations.

1 For an early article on this topic, see Matthew J Bester, 'A Wreck on the Info-Bhan: Electronic Mail and the Destruction of Evidence' (1998) 6 CommLaw Conspectus 75.
2 *Nucleus Information Systems v Palmer* [2003] EWHC 2013 (Ch), where employees used software in an attempt to overwrite the data on computers owned by the company before being returned; *R v Smith (Graham Westgarth), R v Jayson (Mike)* [2002] EWCA Crim 683, [2003] 1 Cr App R 13, [2002] Crim LR 659 in which Jayson deleted a number of abusive images of children that were recovered; *Prest v Marc Rich & Company Investment AG* [2006] EWHC 927 (Comm), 2006 WL 2850945, where it was alleged the claimant deliberately deleted documents on his laptop computer; *R v Porter* [2006] EWCA Crim 560 where it was held that it is a matter for the members of a jury to determine whether files were in the 'possession' of the accused, where the accused placed the files in the recycle bin, and the recycle bin was then deleted – the files were incapable of being recovered (and thus viewed) without the use of specialist forensic techniques and equipment provided by the US Federal Government which was not available to the public; *R v Grout* [2011] EWCA Crim 299, where the day before the appellant's arrest, he re-formatted his computer, so that his computer contained no MSN history of any kind before that date.

9.102 In practice, a person might delete emails and files as a matter of routine, and the organization might fail to realise that it has back-up copies of all the relevant data,[1] or the organization might have back-up data to deal with situations where data is deleted, whether inadvertently or deliberately. For instance, in *Nobel Resources SA v Gross*,[2] Mr Gross attempted to delete SMS messages that might have incriminated him. Several thousand of these messages were recovered from various places: from back-ups of his personal mobile telephone and the BlackBerry of the person to whom the messages were sent. Copies were also found in a back-up file on his laptop computer shortly before trial; they were also on the forensic image of his laptop taken by his forensic experts, and on a CD of his personal files that he only disclosed during the course of the trial. Mrs Justice Gloster DBE said: 'with the assistance of one Jimmy Weston, an IT expert, Mr Gross had deliberately changed the time settings on the laptop to conceal the fact that he himself had made the deletions; and that the last recorded logon time with his user ID reflected this'.[3]

1 As in *Fiona Trust & Holding Corporation v Privalov* [2010] EWHC 3199 (Comm), [1393]; [1397]–[1404].
2 [2009] EWHC 1435 (Comm).
3 [2009] EWHC 1435 (Comm), [54].

9.103 Data destruction adds a great deal of complexity to both civil litigation and the investigation of alleged crimes. On occasions, a party may have a reasonable suspicion that the other party might intend to delete files, or has already deleted files, although the technical issues relating to such allegations can serve to confuse.[1] In *United States of America v Triumph Capital Group, Inc.*,[2] McCarthy, the CEO and controlling shareholder of Triumph, Spadoni, Triumph's Vice President and General Counsel, together with a number of others, were accused of a variety of offences relating to racketeering, including bribery, obstruction of justice and witness tampering. It came to the notice of the US government that Spadoni was alleged to have purchased a software program to purge his computer of incriminating evidence. Triumph was ordered to deliver up the relevant computer for forensic tests. The tests revealed that relevant data had been

deleted, and the deleted files were recovered. A search of the recovered Internet cache files revealed evidence of other offences. This caused the investigator to obtain a further warrant to search and seize evidence of the further crimes. In *L C Services v Brown*[3] the operating system on Brown's computer had been changed or reinstalled at the time the claimants were pursuing disclosure of documents by the defendants, but a digital evidence professional was able to recover the remains of email communications. The recovered evidence was sufficient to incriminate him and he was held liable for breach of fiduciary duties to the plaintiffs, his ex-employer.

1 The decision by the Supreme Court of Delaware in the case of *Genger v TR Investors, LLC*, 26 A.3d 180 (2011), 2011 WL 28028322011, upholding a finding of spoliation by the trial judge, was examined in detail in Daniel B Garrie and Bill Spernow, 'Legally correct but technologically off the mark' (2010) 9 Northwestern Journal of Technology & Intellectual Property 1, in which the authors took the view that the judges failed to understand what had occurred in technical terms.
2 211 F.R.D. 31 (D.Conn. 2002).
3 [2003] EWHC 3024 (QB), [53] and [54].

9.104 Where there is a reasonable suspicion that a party might delete files, as in the proceedings leading up to divorce in the case of *Ranta v Ranta*,[1] it may be possible to obtain an order to prevent a party from deleting, removing or uninstalling any programs, files or folders.[2] Sanctions may follow for deleting files, depending on the seriousness of the action, where a party deliberately wipes hard drives after a court has ordered their production, as in *Electronic Funds Solutions v Murphy*.[3] Furthermore, it is not inconceivable for a court to order a party to search for relevant documents in back-up tapes and archives and to provide information about data that have been deleted.[4]

1 2004 WL 504588 (Conn.Super.).
2 See *Takenaka (UK) Ltd and Corfe v Frankl* [2001] EWCA Civ 348 where patterns of online behaviour were analysed to establish whether it was more likely that defamatory emails were sent to the defendant's wife and used to show that certain pieces of software were used in close proximity to each other and therefore made it more likely that the suspect had sent the emails; *L C Services v Brown* [2003] EWHC 3024 (QB) at [60] and [68]; *Douglas v Hello! Ltd* [2003] EWHC 55 (Ch), [2003] 1 All ER 1087 at [37], [42]–[43] and [87]–[104]; *Crown Dilmun v Sutton* [2004] EWHC 52 (Ch) at [27] and [144]; *LTE Scientific Ltd v Thomas* [2005] EWHC 7 (QB); *Prest v Marc Rich & Company Investment AG* [2006] EWHC 927 (Comm) 2006 WL 2850945 at [8] and [10]; *Sectrack NV v Satamatics Ltd* [2007] EWHC 3003 (Comm) at [6] and [7]; *Noble Resources SA v Gross* [2009] EWHC 1435 (Comm) at [53] and [57]–[58]; *First Conferences Services Ltd v Bracchi* [2009] EWHC 2176 (Ch); note also *Crowson Fabrics Limited v Rider* [2007] EWHC 2942 (Ch); *Rybak v Langbar International Ltd* [2010] EWHC 2015 (Ch). For the USA, see Shira A Scheindlin and Kanchana Wangkeo, 'Electronic discovery sanctions in the twenty-first century' (2004) 11 Mich Telecomm Tech L Rev 71; *Arista Records, L.L.C., v Tschirhart*, 241 F.R.D. 462 (2006), 2006 WL 2728927; Dan H Willoughby Jr, Rose H Jones and Gregory R Antine, '*Sanctions for e-discovery violations: By the numbers*' (2010) 60 Duke Law Journal 789; Charles W Adams, 'Spoliation of electronic evidence: Sanctions versus advocacy' (2011) 8 Mich Telecomm Tech L Rev 1.
3 36 Cal.Rptr.3d 663 (Cal. Ct. App. 2005).
4 *Zhou v Pittsburg State University*, 2003 WL 1905988 (D.Kan.); in relation to digital audio files (including case law), see Alan F Blakley, 'Digital audio files in litigation' (2007) 2 Journal of Legal Technology Risk Management 1.

9.105 As indicated above, the use of these tools has been the result of legal requirements to ensure data security and privacy protection, which means that increasingly, they come with official guarantees that promise that the wiped data cannot be reconstructed by criminals[1] – and as a side effect, the police cannot reconstruct the data either. For instance, to provide legal entities with the assurance that they comply with the law,

such programs typically allow default settings that erase data automatically every time a computer is shut down, or every time someone tries to obtain access to a file without the password. This makes it increasingly problematic to infer criminal intent to hide data when evidence of disc cleaning is found.

1 For instance, Richard Kissel, Andrew Regenscheid, Matthew Scholl and Kevin Stine, *Guidelines for Media Sanitization* (NIST Special Publication 800-88, Revision 1, December 2014), <http://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-88r1.pdf>.

9.106 The physical destruction of a computer, including the hard drive, will ensure the data (without investing in costly reconstruction and recovery services) is lost, as in *Strasser v Yalamanchi*.[1] In this case, it was claimed that a hard drive containing relevant data had been severely damaged by lightning, and an employee saw fit to dispose of the computer as a result. In response to the extensive pre-trial actions and the failure to provide an adequate reason for the destruction of the computer while litigation was under way, the trial judge subsequently instructed the members of the jury that the negligent destruction of evidence may be inferred from the failure of the appellant to preserve and maintain evidence. The appeal court subsequently upheld the decision.

1 783 So.2d 1087 (Fla.App. 4 Dist. 2001).

9.107 Generally speaking, the most secure way to prevent computer forensics is the physical destruction of the hard drive, or short of that, degaussing in a strong magnetic field. Degaussing with an approved degausser is, for some highly sensitive military or national security applications, the required method of data destruction.[1] As discussed above, in comparison to the deliberate attempt at destruction to prevent others from obtaining evidence, paper copy files of underlying source documents may be destroyed for perfectly legitimate reasons, and reliance might subsequently be made on the version held in electronic form. This tends to occur when organizations attempt to reduce the cost of storage of paper documents but fail to consider the cost of electronic storage and the need to deal with old data when a system is upgraded. In the case of *Heveafil Sdn. Bdh., v United States*,[2] the US Department of Commerce refused to accept a copy of a database containing a bill of materials stored on a computer diskette as a means of verifying the cost information in an investigation into anti-dumping extruded rubber. Heveafil claimed that the database held on the diskette had been taken from the mainframe, it used the previous version in the course of normal business, and asserted that the database on the diskette contained an exact duplicate of the database developed on the mainframe computer. In an appeal from the US Court of International Trade, the Court of Appeals for the Federal Circuit accepted the argument by the Department of Commerce that it could reject the data on the diskette as not having been properly authenticated, and a finding of adverse inference was admissible in the circumstances. The assertions by Heveafil were not sufficient, because it failed to provide evidence of the veracity of the contents of the diskette, such as explanations of how the copy was made. The company merely copied data from the mainframe and then deleted the first in time data as well as the underlying paper versions. In doing so, they failed to provide a trail of evidence to demonstrate the procedures undertaken to provide for the veracity of the copy.

1 <www.nsa.gov/ia/_files/government/MDG/EPL_Degausser25June2012.pdf>.
2 58 Fed.Appx. 843; 2003 WL 1466193 (Fed.Cir.); 25 ITRD 1128.

9.108 As mentioned above, the social context can be crucial in interpreting electronic evidence. While clicking on the delete icon, as we have argued, is not a way actually to destroy evidence, and can furthermore be seen as an intentional attempt to destroy evidence and pervert the course of justice, the opposite question also arises: under what circumstances can the law interpret a user's failed attempt to destroy a file as a sign that he wanted to rid himself of possession of an illegal item? An innocent user who accidentally downloads an illegal picture, or finds one on a second-hand computer, may think that by deleting the item he has successfully rid himself of it. The law on possession of illegal material may or may not take the same view, if, for an average user, it is very easy to recover the item in question, and it is thus possible to use the 'paper bin' as a convenient hidden storage space.

9.109 All the methods of data deletion described above have been developed for data stored on traditional magnetic media. But increasingly, new storage media look set to challenge anti-forensic measures and also thwart the efforts of investigators. With traditional magnetic storage media, 'bad sectors' can create inaccessible parts of the hard drive that are 'accidentally' protected from many cleaning utilities. Solid-state drives (SSDs), unlike traditional magnetic discs such as hard disk drives, do not have any moving mechanical components but use integrated circuits to store data persistently. SSDs pose new problems for the recovery of data, because they store data in ways that are much more non-linear and complex than that of traditional hard disk drives.[1] However, programs such as *Parted Magic* claim to provide safe data cleaning for SSDs.

1 Bell and Boddington, 'Solid state drives: The beginning of the end for current practice in digital forensic recovery?'.

9.110 Several new filing systems increase data permanence either by design – to prevent accidental data loss – or by accident. For instance, journaling file systems record write operations in a number of different locations, which means data 'leftovers' may exist in places 'outside' the nominal file storage location. RAID and anti-fragmentation techniques may also result in file data being written to multiple locations. In SSDs, for instance, if the same part of the drive is written over and over again, this will have the effect of 'wearing it out' prematurely. To counteract that, technologies are built into SSDs called 'wear levelling', which relocates blocks of data between the time when they are originally written and the time when they are overwritten. This has the effect of preventing the erasure of data.

9.111 From a legal and evidential perspective, it is necessary to have some knowledge of the differences these storage media entail for data deletion and data retrieval to interpret the findings of the digital evidence professional correctly. The easier it is to securely delete data with off-the-shelf, easily customisable tools, the less convincing the inference to an intentional attempt to hide evidence is. Finding evidence for the deletion of data from traditional hard drives is therefore different from evidence of deletion from new and more advanced storage systems, where data erasure requires specialist knowledge and considerable efforts.

9.112 The question that remains is what inferences, if any, can be drawn from the *absence* of evidence if data have been successfully deleted. A defence lawyer may want to argue that according to the prosecution story, some traces of illegal activity *ought*

to have been found on his client's computer, using the *absence* of such evidence as an argument to undermine the prosecution case. How convincing the argument is may well depend on the type of storage medium used and the nature of file systems deployed. As noted with wear levelling, there are also increasingly automated 'housekeeping operations' being carried out by computers on files. In the past, finding that an illegal file, say of images of child sexual abuse, had been moved and copied to several places of a hard drive would have been evidence that the suspect knew of, and knowingly handled, the file in question. Increasingly, this inference depends on the storage medium, and if a number of copies at different parts of the drive existed, the possibility that these could have been the result of automated actions by the computer. Finally, for several legal purposes, a party may have to prove that it either took all reasonable steps to delete certain files, for instance in an action for damages after a data security breach, or that it took every reasonable effort to produce data, for instance, in response to a court order as part of the disclosure or discovery process. The type of evidence required to document that all reasonable steps were taken to either securely delete the data, or to recover lost data, will depend on the precise nature of the storage medium.

9.113 A separate way of destroying data at the filesystem level is by the deletion of filesystem-wide encryption keys. Mobile telephones and several desktop operating systems increasingly feature encrypted filesystems that use a private key for unlocking the data in the filesystem. This key has to be unlocked and made available for en- and de-cryption each time the computer or telephone is booted, turned on after a longer delay, or after the key memory-retention period has expired. Data that is written to the persistent filesystem is encrypted using a unique (system specific, locally generated) private key, which is then secured (and unlocked upon demand) using a PIN, swipe pattern or fingerprint. Upon unlocking the telephone, this key is decrypted to enable full access. Destroying the private key, however, makes it virtually impossible to retrieve the data on the telephone, provided the cryptography and the implementation of this feature is done to exacting security standards. A modern smartphone may then destroy all data if a certain number of attempts are made to unlock the private key with a wrong or false finger print or access code.[1]

1 A good example is the implementation of this system in iOS for Apple smartphones. It is described extensively in the iOS 9.3 or later security guide (May 2016), available at <www.apple.com/business/docs/iOS_Security_Guide.pdf>.

Falsifying data

9.114 Tampering with electronic evidence is not new. An early example of erasing part of a tape recording and re-recording part of a conversation occurred in the UK in 1955.[1] In *R v Sinha*,[2] medical data recorded on a computer was altered after the death of a patient, giving rise to a charge of perverting the course of justice. In the recent case of *Freemont (Denbigh) Ltd v Knight Frank LLP*,[3] one witness concocted evidence by creating documents in the form of a series of notes of discussions, which included statements which had not been made during the course of the discussions,[4] and to avoid detection, had the hard drives of older computers destroyed when the firm upgraded its computer systems.[5] Attempts to adduce fraudulent evidence before a court are rare, but increasing.[6] For instance, Bruce Hyman, who had been a prominent British television and radio producer before qualifying as a barrister later in life, created a false judgment for a friend. His deception was uncovered and he was subsequently

convicted for perverting the course of justice and sentenced to a term of imprisonment of twelve months and ordered to pay £3,000 to his victim in compensation and Crown expenses of £3,745 – the first barrister to be so convicted, and he was subsequently disbarred by the Bar Standards Board.[7] In another case in Japan, a prosecutor altered electronic evidence in a case he was investigating, and was subsequently convicted and imprisoned for 18 months.[8]

1 'Recording as testimony to truth' [1955] Crim LR 2; [1954] SJ 98, 794.
2 [1995] Crim LR 68, CA.
3 [2014] EWHC 3347 (Ch).
4 Although the judge did not have to determine precisely how the evidence was concocted, and he considered the possibility of amended computer files, at [56], he concluded on other evidence that the evidence was concocted, for which see [116], [123] and [140].
5 At [56]–[60].
6 *Premier Homes and Land Corporation v Cheswell, Inc.*, 240 F.Supp.2d 97 (D.Mass. 2002) for fabrication of an email; *People v Superior Court of Sacramento County*, 2004 WL 1468698 (Cal.App. 3 Dist.) for fabrication of letters on a computer after the event; *ISTIL Group Inc v Zahoor* [2003] EWHC 165 (Ch), [2003] All ER (D) 210 (Feb) at [106] and [111] for a forged document; *Fiona Trust & Holding Corporation v Privalov* [2010] EWHC 3199 (Comm) at [1405]–[1430] for a forged and back-dated agreement and employment contract; for forged emails, *Apex Global Management Ltd v FI Call Ltd* [2015] EWHC 3269 (Ch); in a criminal context, see *R v Brooker* [2014] EWCA Crim 1998 (available in the LexisNexis electronic database), where Brooker sent text messages from a second mobile telephone in her possession, claiming that her boyfriend sent them.
7 Angella Johnson, 'How my barrister forged evidence against my husband – and now faces jail', *The Mail* (8 September 2007); Steven Morris, 'Barrister becomes first to be jailed for perverting justice', *The Guardian* (20 September 2007); Simon de Bruxelles, 'Barrister jailed for trying to frame man with fake e-mail', *Timesonline* (20 September 2007).
8 Hironao Kaneko, case translation and commentary in 'Heisei 22 Nen (Wa) 5356 Gou' (2012) 9 Digital Evidence and Electronic Signature Law Review 109.

9.115 However, it is conceivable, given the ease with which electronic data is so easily manipulated and altered, that attempts will be made in the future to falsify and alter documents even before a trial ever takes place, or to create vast swathes of 'evidence' of a complete set of legal proceedings. This happened in *Islamic Investment Company of the Gulf (Bahamas) Ltd v Symphony Gems NV*.[3] As explained in the judgment of Hamblen J:

> From the end of October 2010 until December 2013 [the lawyer] conducted fictitious litigation for RM. That litigation involved fictitious hearings before the Commercial Court and the Court of Appeal; purported judgments of those courts; purported sealed court orders; a purported hearing transcript; purported skeleton arguments; purported correspondence with court officials and the Claimant's solicitors, Norton Rose; the fictitious instruction and engagement of various counsel, and telephone conferences involving the impersonation of his senior partner and of leading counsel. None of this reflected reality. Throughout that period there was in fact no contact with Norton Rose or the court.[2]

153 [2014] EWHC 3777 (Comm).
154 [2014] EWHC 3777 (Comm), [4].

9.116 Even such mundane matters such as proof of parking violations have been subject to the alteration of electronic evidence. In the case of Kevin Maguire, he had parked his car in Market Place in Bury town centre, Greater Manchester at 7.15am on 31 August 2003. He returned at 5pm to find he had been given a parking ticket at 9.15am. Normally there were no restrictions on a Sunday, and when he parked his car,

there were no signs to indicate there were any temporary restrictions in place. There were no signs because the NCP staff did not put them up on the previous night as there was a high likelihood that the signs could be pulled down or damaged by revellers overnight. In fact, the signs were put up after Mr Maguire had parked his car. When Mr Maguire complained to the NCP, it was asserted that he had parked illegally and he was sent a photograph of his parked car, which was dated 30 August 2003. Mr Maguire appealed against the parking fine. It transpired that one Gavin Moses, a member of the NCP staff, had altered the date on the digital photograph from 31 August to 30 August, so that it appeared that Mr Maguire had parked illegally. Mr Maguire was cleared of illegal parking and was awarded costs. Gavin Moses subsequently entered a plea of guilty when he was prosecuted for perverting the course of justice, and was sentenced to 150 hours of community service.[1]

1 '"Fit up" parking warden sentenced', BBC News (28 January 2005) <http://news.bbc.co.uk/1/hi/england/manchester/4216539.stm>. A further article was published by a Manchester website dated 27 January 2005, but the webpage is no longer active.

9.117 In Singapore, Ruddy Lim altered the monthly salary on his payslip from DLA Piper Singapore Pte Ltd to read $65,000, rather than $25,000. The description of his method is set out in the judgment:

> The Accused testified that he first created Exhibit P2 in his laptop computer some time between 12 and 14 November 2006. He was travelling in Jakarta at the time, and carried a soft copy of the DLA Piper logo in his laptop for preparing marketing materials. He created a document in the word-processing programme, Word, by typing out the text and numbers of the false payslip. He cut and pasted the DLA Piper logo onto the Word document. He then copied the image of the company stamp (with the office manager's signature) from his original payslip … using software from Adobe, and electronically affixed the image onto the Word document. During this time, the Word document existed only in soft copy. When he returned to Singapore, he printed out the Word document on 14 November 2006, then scanned it into the Xerox machine so that a 'pdf' version of the false payslip would be created. He wanted to convert it from Word document format into 'pdf' because the former was 'editable', while the latter was a 'fixed format'. He then emailed the resulting document … to [his prospective employer].[1]

He was found guilty of forgery and sentenced to two months' imprisonment.[2]

1 *Public Prosecutor v Rudy Lim* [2010] SGDC 174, [17].

2 According to the conclusions on page 56 of *Report of Digital Forensic Analysis* (26 March 2012) by Stroz Friedberg and submitted as evidence in the case of *Paul D. Ceglia v Mark Zuckerberg, Individually, and Facebook, Inc.*, Civil Action No: 10-cv-00569-RJA, Stroz Friedberg determined that it had 'found direct and compelling digital forensic evidence that the documents relied upon by Mr. Ceglia to support his claim were forged.' Available at <http://cdn.arstechnica.net/wp-content/uploads/2014/08/strozreport.pdf>.

9.118 Considerably more attention will have to be paid to demonstrate the integrity of electronic data in the future, which in turn will help substantiate the claim for authenticity to reflect the reliability of the data. In all of these cases, the changes to the data were carried out manually. Anti-computer forensics increasingly provide tools to alter data, and in particular the crucial metadata, automatically, thus diminishing the evidential value of the data that can be recovered. 'Backtrack' or 'Transmogrify', for instance, can change the extension of files by turning .exe (application) files into .docx

(Word document) files, thereby hiding their malicious character. 'Timestomp' can change the timestamps of files, the metadata that records the creation and alteration of a file.[1] Randomisers can automatically generate random file names, and criminals can use tools that replace Roman letters with identical-looking Cyrillic ones. Both approaches defeat data-mining techniques that look for 'known bad files' or signatures of known illegal images. Many of these tools were developed by software developers, who wanted to test the reliability of common forensic tools such as Encase. Vincent Lui, one of the most prolific developers of tools with anti-forensic implications, concludes that the 'unfortunate truth' is that the presumption of reliability is 'unjustified' and the justice system is 'not sufficiently sceptical of that which is offered up as proof'.[2]

1 Hamid Jahankhani and Elidon Beqiri, 'Digital evidence manipulation using anti-forensic tools and techniques' in Hamid Jahankhani, David Lilburn Watson and Gianluigi Me (eds), *Handbook of Electronic Security and Digital Forensics* (World Scientific Publishing Co Pte Ltd 2010).
2 Eric Van Buskirk and Vincent T Liu, 'Digital evidence: Challenging the presumption of reliability' (2006) 1 Journal of Digital Forensic Practice 18, 25.

9.119 Other tools have legitimate objectives such as privacy protection. For instance, to prevent companies from data mining our behaviour when using a search engine, software can be used to create a large number of chance queries to create random noise.[1] Since a record of keyword searches can also have evidential value in a criminal trial, to establish the interest of the suspect in certain poisons or drugs, these tools can also cast doubt on the reliability of the log data that documents the searches carried out on a suspect's computer. Since the search terms had been automatically generated, any inference that the user of the machine intentionally searched for a specific term becomes problematic.[2]

1 Ye Shaozhi, Felix Wu, Raju Pandey and Hao Chen, *Noise Injection for Search Privacy Protection* (2009) <http://escholarship.org/uc/item/08k1004m>.
2 This is clearly illustrated in the case of *State of Connecticut v Julie Amero* (Docket number CR-04-93292; Superior Court, New London Judicial District at Norwich, GA 21; 3, 4 and 5 January 2007). For a detailed analysis of this case, see Mason, *International Electronic Evidence*, xxxvi–lxxv.

Hiding data

9.120 Tampering with and destroying data works best when the criminal no longer needs the data. For possession crimes such as the possession of illegal images, this is not possible. Hiding the data rather than destroying or altering it therefore becomes an important objective. Cryptography is the best known anti-forensic method to hide data from third parties. Due to its importance as a dual use technology with important roles for privacy and data security, and also because of the complex legal issues involved with cryptography, this is considered in the chapter on encrypted data.

9.121 Another well-known method of hiding data is steganography. Steganography is the method of hiding a message inside a digital object, which may be a graphic, a picture, a film or a sound clip. The sender is able to hide a message in a seemingly innocuous file, and the recipient can retrieve the message upon receipt. Other methods used to hide data include writing data to slack space or space that has not been allocated for use, hiding data on a hard drive in a secret partition, and the transmission of data under the cover of transmission protocols. Various types of commercial and free software are available to perform steganography on data. It can be relatively

difficult to detect hidden data within a file, and the communication can be even more difficult to uncover if the message has been compressed and encrypted before being hidden in the carrier. At present, it is unlikely that many investigators will undertake a routine examination for hidden data.[1]

1 Brent T McBride, Gilbert L Peterson and Steven C Gustafson, 'A new blind method for detecting novel steganography' (2005) 2 Digital Investigation 50; a wide range of references on this topic is provided in Gary C Kessler, 'An overview of steganography for the computer forensics examiner' (2004) 6 Forensic Science Communications – for an update of this article to February 2015, see <www.garykessler.net/library/fsc_stego.html>; Rachel Zax and Frank Adelstein, 'FAUST: Forensic artifacts of uninstalled steganography tools' (2009) 6 Digital Investigation 25.

9.122 There are now various tools available that facilitate the hiding of data in places on the hard drive that are less likely to be inspected. In this sense, they are the mirror images of the deletion tools discussed above. Deletion tools aim to securely delete any trace of an incriminating file, regardless of where on the computer a copy may be hiding. Conversely, 'Slacker' breaks up a file and stores individual pieces of it in the slack space left at the end of files, making it look like random noise to forensic tools – imagine just two digits each of a stolen credit card number stored in the unused part of a legitimate file. Slacker then enables the data to be reassembled as required.[1] One of the problems with these tools is that they develop faster than it is possible to train digital evidence professionals, and even more importantly, faster than the development of sound, tested and agreed standards. This not only makes the detection of evidence more difficult, it also raises issues about the admissibility of forensic expertise.

1 Hal Berghel, 'Hiding data, forensics, and anti-forensics' (2007) 50 Communications of the ACM 15.

Attacks against computer forensics

9.123 Arguably, the latest addition to the inventory of anti-computer forensics is attacks against the investigator and her tools. As noted above, digital forensics is highly dependent on software tools. To create evidence that is admissible, these tools have to be evaluated and tested, and the results ideally published in openly available, peer-reviewed scientific publications. Indeed, some of the most popular tools are open source: that is, their source code is freely available. One of the benefits of this approach is not only a high degree of transparency when it comes to assessing the reliability of data generated by these tools, but also the ability for security professionals to improve them and to adapt them to local situations.[1] However, it also enables criminals to develop tools that interfere directly with the evidence collection process and infiltrate the software that tries to analyse a suspect's computer. This can either be done by undermining the integrity of the data that is collected, for instance by changing the hash value of the bit copy that the software creates (thus violating the continuity of evidence by casting reasonable doubt on the authenticity of the copy) or by forcing the analysis tool to either overlook incriminating data, or to report misleading information about it,[2] which indicates that it cannot be right to equate such a tool with, say, a photocopier.[3] The conflict between admissibility standards such as Daubert in the US that rely on publicly available information about forensic techniques and the need to protect the integrity of the analysis tools will be difficult to bridge.

1 Erin E Kenneally, 'Gatekeeping Out of the Box: Open Source Software as a Mechanism to Assess the Reliability of Digital Evidence' (2001) 6 Va JL & Tech 13.

2 Chris K Ridder, 'Evidentiary implications of potential security weaknesses in forensic software' (2009) 1 International Journal of Digital Crime and Forensics 80.
3 *Williford v State of Texas*, 127 S.W.3d 309 (Tex.App.—Eastland 2004).

Trail obfuscation

9.124 Trail obfuscation combines the deliberate attempt at tampering, deleting and hiding data with the taking of measures to frustrate investigations, conceal identities and evade enforcement actions.[1] In many investigations, the data held on the suspect's computer is only one part of the prosecution's case. The other, equally important, set of data will come from the Internet and relate to the suspect's browsing behaviour, or the computer of his victim in the case of a hacking offence: the origin of the data, the websites he visited, and the activities he undertook. Obfuscating the trail that such activities leave behind on the Internet is therefore an important aspect of anti-computer forensics. It includes various anonymity-protection tools such as VPNs or anonymous remailers to hide browsing activity, or the use of spoofed or zombified accounts when sending malicious emails or spam, or the launch of a denial of service attack. 'Zombified accounts', as discussed in more detail below, demonstrate a specific side effect of anti-computer forensics. One way for a criminal to hide his activities is to take over the computer of a third party, for instance, after inserting a Trojan horse program, discussed in more detail below, and using this third party machine to carry out illegal activities. This not only hides the true perpetrator from the investigators, it also creates data that can falsely incriminate an innocent party.[2]

1 In the civil context, see *EMI Records Ltd v British Sky Broadcasting Ltd* [2013] Bus LR 884, [2013] WLR(D) 86, [2013] EWHC 379 (Ch).
2 Srinivas Mukkamala and Andrew H Sung, 'Identifying significant features for network forensic analysis using artificial intelligence techniques' (2003) 1 Intl J of Digital Evidence; Bruce J Nikkel, 'Domain name forensics: A systematic approach to investigating an internet presence' (2004) 1 Digital Investigation 247; Bruce J Nikkel, 'Improving evidence acquisition from live network sources' (2006) 3 Digital Investigation 89; Eoghan Casey and Aaron Stanley, 'Tool review – remote forensic preservation and examination tools' (2004) 1 Digital Investigation 284; Omer Demir, Ping Ji and Jinwoo Kim, 'Packet marking and auditing for network forensics' (2007) 6 Intl J of Digital Evidence.

9.125 The range of tasks performed by such malicious software is probably only restricted by the imagination of the person who creates the program. A number of cases in the criminal courts where people have been accused of being in possession of abusive images of children on their computers have used the defence that some form of malicious software caused data to be downloaded to their computers or enabled a third party to obtain access to their computers without the permission of the computers' owners.[1] In the case of *R v Caffrey*,[2] the defendant was charged with causing unauthorized modification of computer material under s 3(1) of the Computer Misuse Act 1990. The prosecution alleged that the defendant sent a deluge of electronic data from his computer to a computer server operated in the Port of Houston, Texas, USA, the effect of which was to cause the computer at the Port of Houston to shut down. He claimed, in his defence, that unknown hackers obtained control of his computer and then launched a number of programs to attack the computer at the Port of Houston. The forensic examiner for the prosecution could not find any evidence of a Trojan horse on the computer. The defence claimed that it was impossible for every file to have been tested, and that the Trojan horse file might have had a facility to destroy itself, leaving no traces of having resided on his computer. The forensic examiner for

the prosecution disputed that, stating that a Trojan horse would leave a trace on the computer. The jury acquitted Mr Caffrey.[3]

1 *R v Schofield* (April 2003, unreported), Reading Crown Court, and *R v Green* (October 2003, unreported), Exeter Crown Court.
2 (October 2003, unreported), Southwark Crown Court.
3 Esther George, 'Casenote' (2004) 1 Digital Investigation 89; Susan Brenner, Brian Carrier and Jef Henninger, 'The Trojan Horse defense in cybercrime cases' (2004) 21 Santa Clara High Tech LJ 1; the first Trojan horse case in the People' Republic of China was prosecuted in 2009: Jihong Chen, 'The first "Trojan Horse" case prosecuted in China' (2010) 7 Digital Evidence and Electronic Signature Law Review 107; Alex Xia and Julia Peng, 'First "Trojan horse" case prosecuted for illegal invasion of computer systems in China' (2009) 25 Computer Law & Security Review 298.

9.126 It should be noted that just because an individual may have such materials on his computer, it does not follow that he was responsible for downloading them onto his computer. It is important for any digital evidence professional to report his findings within the context of what the technology is capable of doing. For instance, it is possible to introduce malicious software through web pages without the permission of the website owner. When a person visits a website, software could redirect the computer to undesirable websites, and the computer will automatically download unwanted material onto the temporary cache file of the computer without the user's permission or knowledge.[1]

1 For which, see Daniel Bilar, 'Known knowns, known unknowns and unknown unknowns: anti-virus issues, malicious software and internet attacks for non-technical audiences' (2009) 6 Digital Evidence and Electronic Signature Law Review 123 (in which the author illustrates the ease by which third parties can obtain control of computers without the authority of the owner or user); Megan Carney and Marc Rogers, 'The Trojan made me do it: A first step in statistical based computer forensics event reconstruction' (2004) 2 Intl J of Digital Evidence.

9.127 A *Trojan horse* is a malicious software program containing hidden code that is designed to conceal itself in a computer as if it were legitimate software. When activated, the software will perform an operation that is not authorized by the user, such as the destruction of data (including the entire hard drive), the collection of data on a computer and transmission to a third party without the user being aware of what is happening, the counteraction of security measures installed on a computer, and the instruction of the computer to perform tasks such as to take part in a denial of service attack, or permit the creator of the program to obtain access to the computer. Just like the other large group of malware, viruses, Trojan horses pose a Janus face for computer forensics. Finding a virus or a Trojan infection can be direct evidence for possible charges under the Computer Misuse Act 1990 as amounting to an unauthorized modification of computer systems. At the same time, this can also be indirect evidence that the computer at the centre of an investigation has been tampered with and that the crime scene is 'contaminated'.

9.128 The dual use nature of many of the tools used for anti-computer forensics has been noted above. On the one hand, these tools protect our privacy against criminals, but they also protect the privacy of criminals from police investigations. A similar analysis applies to spyware such as Trojan horses. On the one hand, they allow criminals to obtain access to credit card details or passwords. On the other, they have the potential to allow the police to obtain access to the activities of criminals – that is, if the police succeed in planting such a program on the suspect's computer. Attempts to use

malware for investigative purposes have caused legal controversy in some countries. In Germany, the Constitutional Court ruled against such clandestine surveillance after prosecutors applied for warrants to permit their use. In the discussion before the court, evidence was also given from computer specialists about the security and evidential implications of these 'Federal Trojans'. To work efficiently, they must not be detected by commercial anti-virus software. This can be achieved either by the tacit collaboration of the anti-virus software vendors, or by using the ingenuity of programmers employed by the police. In either case, the result will be malware that cannot be easily detected. One obvious danger is that criminals can get hold of and in turn hijack the code for this 'official' malware once it was planted on their machines, which would give them in effect a 'master key' for all computer systems. In such an event, it would become much easier for the defence to mount 'Caffrey style' arguments, and all computers could become crime scenes with compromised integrity.[1]

1 Wiebke Abel and Burkhard Schafer, 'The German Constitutional Court on the right in confidentiality and integrity of information technology systems – a case report on BVerfG, NJW 2008, 822' (2009) 6 SCRIPT-ed 106.

9.129 A final complication is created by the desire to protect users against malware. The use of Trojans lies at the heart of distributed denial of service attacks, a significant threat to the functioning of the Internet. Preventing malware has therefore become a high priority for police and commerce. Ordinary computer users, who often fail to take appropriate steps to protect their computers against interference by criminals, are the weakest link. The Trusted Computing Initiative is one possible answer to this problem. It would allow a coalition of software and hardware developers much more direct access to computers, ensuring that all their defence mechanisms work as specified, and that no unauthorized program is run on them. While this approach is promising in its potential to reduce computer criminality, for the interpretation of electronic evidence, it carries several challenges. Since computer forensic tools too are essentially a form of 'spyware', common forensic applications may not work any longer in a trusted computing environment. Even worse, the philosophy of trusted computing is premised on belief that to protect the user against criminal activities, the security and control of the computer is improved if it is not just determined by the user but also by organizations. This means that the number of people and organizations that at any given time would have access to users' computers and the data held therein would increase considerably, especially if the keys to users' computers and their devices are compromised. This could in turn cast doubt on the reliability and authenticity of the data found on a computer during a criminal investigation. At the moment, lawyers assume, often naively, that data found on a suspect's computer must have been put there by the person in physical control of the machine (typically, the owner); this inference would look increasingly doubtful in a trusted computer environment.[1]

1 Yianna Danidou and Burkhard Schafer, 'Trusted computing and the digital crime scene' (2011) 8 Digital Evidence and Electronic Signature Law Review 111.

Conclusions and future considerations

9.130 Electronic evidence has been with us for a long time. The widespread use of the Internet, mobile telephones and smartphones means that most lawyers now have

to deal with electronic evidence.[1] This can be straight forward only in simple cases, but not otherwise where the parties challenge the data. For this reason, lawyers must familiarise themselves with electronic evidence and understand the need to scrutinise the qualifications and conclusions of digital evidence professionals.

1 Graeme Horsman and Lynne R Conniss, 'Investigating evidence of mobile phone usage by drivers in road traffic accidents' (2015) 12 Digital Investigation S30.

9.131 One of the major difficulties in investigating evidence in digital form also relates to the incompatibility of formats used to store digital data. The problems arise when an investigator has to deal with different disk image formats. The difficulty is compounded when dealing with different types of electronic evidence, such as network data logs, or the contents of mobile devices.[1] Cloud computing and trusted computing affect the way digital evidence professionals obtain evidence, which means that great care must be taken over how such evidence is obtained, which will doubtless be the subject of careful cross-examination.[2] In addition, the methods used by attackers in the digital environment will mean it is increasingly necessary to take into consideration the use of rarer techniques to obtain evidence in the future.[3]

1 Barrie Mellars, 'Forensic examination of mobile phones' (2004) 1 Digital Investigation 266; Adam Laurie, 'Digital detective – Bluetooth' (2006) 3 Digital Investigation 17; Rick Ayers, Sam Brothers and Wayne Jansen, *Guidelines on Mobile Forensics*, Computer Security Division, Information Technology Laboratory, National Institute of Standards and Technology, Special Publication 800-101 Revision (1), Sponsored by the Department of Homeland Security (May 2014) <http://csrc.nist.gov/publications/PubsSPs.html#800-101>.

2 Stephen Mason, 'Trusted computing and forensic investigations' (2005) 2 Digital Investigation 189: this article is merely an introduction to the topic that includes relevant references, and see also 'Trusting your computer to be trusted' (2005) 7 Computer Fraud & Security, with a number of additional references by the same author; an outline of cloud computing that includes relevant references can be found in Stephen Mason and Esther George, 'Digital evidence and "cloud" computing' (2011) 27 Computer Law & Security Review 524; see also a thesis in partial fulfilment of the requirements for the degree of Masters in Forensic Information Technology submitted to the graduate faculty of Computing and Mathematical Sciences at Auckland University of Technology by Michael E Spence, 'Factors influencing digital evidence transfer across international borders: A case study' (2010) <http://aut.researchgateway.ac.nz/handle/10292/1187>; Ian Walden, *Law Enforcement Access in a Cloud Environment* (Legal Studies Research Paper No 74/2011) <http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1781067>; Giuseppe Vaciago, 'Remote forensics and cloud computing: An Italian and European legal overview' (2011) 8 Digital Evidence and Electronic Signature Law Review 124; Josiah Dykstra and Alan T Sherman, 'Acquiring forensic evidence from infrastructure-as-a-service cloud computing: Exploring and evaluating tools, trust, and techniques' (2012) 9 Digital Investigation S90.

3 Kris Harms, 'Forensic analysis of System Restore points in Microsoft Windows XP' (2006) 3 Digital Investigation 151.

9.132 In response to these developments, anti-computer forensics has emerged over the last decade as a significant challenge to the investigation of crimes involving the use of computers and computer-like devices. The arms race between criminals and investigators on the one hand, and the dual use nature of the tools that permit and prevent digital investigations on the other, have created a highly complex interaction that requires careful reflection on the nature of electronic evidence in any individual case, a reflection that has to be constantly updated as new tools emerge. Some of the challenges and conflicts this creates for the process of collecting, evaluating and examining electronic evidence in a legal setting will be difficult to resolve in the near future.

10

Competence of witnesses

Stephen Mason

The need for witnesses

10.1 Concern is sometimes expressed over the competence, knowledge and qualifications of the witness giving evidence as to the trustworthiness of digital data as evidence. In *Wood (Stanley William)*, the Lord Chief Justice explained this as follows:

> This computer was rightly described as a calculating tool. It did not contribute its own knowledge. It merely did a sophisticated calculation which could have been done manually by the chemist and was in fact done by the chemists using the computer programmed by Mr. Kellie whom the Crown called as a witness. The fact that the efficiency of a device is dependent on more than one person does not make any difference in kind. Virtually every device will involve the persons who made it, the persons who calibrated, programmed or set it up (for example with a clock the person who set it to the right time in the first place) and the person who uses or observes the device. In each particular case how many of these people it is appropriate to call must depend on the facts of, and the issues raised and concessions made in that case.[1]

1 (1983) 76 Cr App R 23, 27.

10.2 The complexity of a computer, whatever the nature of the device (whether a hand-held device or a mainframe computer), will give rise to issues of authentication, but a wider range of challenges may also be raised:

> 1. There may be a question about the accuracy or otherwise of the human input. Where the accuracy of the information is challenged, two factors will be pertinent: whether the human beings responsible for inputting the information entered the correct information; and, regardless of the conclusions reached in answering the first point, whether the software harboured an error or a malicious code that acted to change the information that was entered by humans. In the first instance, evidence from those that were responsible for entering the data, if they can be found, will need to be called. In the second instance, the evidence of a suitably knowledgeable digital evidence professional or a suitable technician that is highly familiar with the system will be necessary.
>
> 2. The 'reliability' of the underlying operating system and application software may be at issue. This is a separate question to the first type of challenge, and will require a witness with different skills to the witnesses required in the first example. Here, it may be necessary to call the manufacturer of the hardware, or the developer of the operating system or application, or failing that, an expert in the specific operating or application software.
>
> 3. The mechanisms developed to ensure a system operates properly and efficiently may be at issue. A good example is that of bank ATMs. It is a notorious fact that attacks on ATMs are successful without the use of the card issued to the customer. Because these systems are subject to outward facing threats, the range of experts will be wider when challenges of this nature are made, and will

Stephen Mason, 'Competence of witnesses', in Stephen Mason and Daniel Seng (eds.), *Electronic Evidence* (4th edn, University of London 2017) 339–49.

include experts who work in a bank as well as experts that are familiar with the weaknesses of bank ATM systems.

10.3 The precise nature of the evidence to be given will be governed by the nature of the challenge by the defence in any one case. The observations made by the Lord Chief Justice in *Wood (Stanley William)* were later elaborated by Steyn J, as he then was, in *Minors*, specifically including an observation underlying the rationale for admitting such evidence without adding to the burden of the prosecution:

> The law of evidence must be adapted to the realities of contemporary business practice. Mainframe computers, minicomputers and microcomputers play a pervasive role in our society. Often the only record of a transaction, which nobody can be expected to remember, will be in the memory of a computer. The versatility, power and frequency of use of computers will increase. If computer output cannot relatively readily be used as evidence in criminal cases, much crime (and notably offences involving dishonesty) will in practice be immune from prosecution. On the other hand, computers are not infallible. They do occasionally malfunction. Software systems often have 'bugs'. Unauthorised alteration of information stored on a computer is possible. The phenomenon of a 'virus' attacking computer systems is also well established. Realistically, therefore, computers must be regarded as imperfect devices.[1]

1 *R v Minors (Craig); R v Harper (Giselle Gaile)* [1989] 1 WLR 441, 443.

Separating data reliability from computer reliability

10.4 In the case of *Minors*, the appellant tendered a passbook with false entries purporting to show there was more money held in the account than the £1 that was actually recorded. An auditor, a member of the audit investigation department of the Alliance and Leicester Building Society, who had 14 years' relevant experience and regularly worked with the particular computer, produced the computer record of the complete history of the appellant's account. The last four (forged) entries in the account book were not recorded in the computer print-out. The evidence of the computer print-out was relevant to the question whether there was, in fact, only a balance of £1 in the account. For technical reasons that no longer apply, it was held that the evidence of the building society auditor was wrongly admitted under the provisions of the Police and Criminal Evidence Act 1984 that prevailed at the time.

10.5 In this case, it is pertinent to note that the auditor was properly qualified to testify as to the 'reliability' of the computer. However, it is suggested that the 'reliability' of the computer was not in issue in this case. The issue was whether the information entered into the computer was accurate, and if so, how the accuracy or otherwise of the information could be proved. The 'reliability' of the computer was a separate issue. All the auditor would be doing in such circumstances was to provide evidence as to how the information was transcribed from the passbook to the computer, and whether the methods used by the building society were capable of providing the assurance that the information was accurate.

10.6 In the case of *Harper*,[1] it was alleged that the appellant presented a stolen Capital Card when travelling on a London Transport bus. The relevant sequence of events were as follows. In February 1985 a batch of cards were stolen at Alexandra

Palace railway station; appropriate entries were made by an employee in the 'Lost Book' at the station; the relevant entries were transferred to a computer belonging to British Rail at King's Cross railway station, and the entries were further transferred from this computer to a computer at Waterloo railway station owned by London Regional Transport. At trial, the prosecution relied on a computer print-out from the final computer to show that the card was stolen. The print-out was produced by a revenue protection official who worked at Baker Street station. The judge admitted the evidence, but it was held on appeal that it was incorrect to admit the evidence, because the witness could not, from her own knowledge, testify to the 'reliability' of the computer, and also that the requirements of s 68 of the Police and Criminal Evidence Act 1984 had not been satisfied.[2]

1 *R v Minors (Craig); R v Harper (Giselle Gaile)* [1989] 1 WLR 441; [1989] Crim LR 360.
2 Section 68 of the Police and Criminal Evidence Act 1984 was repealed by the Criminal Justice Act 1988, Schedule 16.

10.7 This decision must be right. However, it is suggested that the 'reliability' of the computer was not relevant given this set of facts. The fatal problem in this instance was a break in the continuity of evidence, because the 'Lost Book' held at Alexandra Palace railway station was missing at the time of the trial. The witness may have been competent to give evidence of the procedures used to register and disseminate the knowledge of the loss of Capital Cards. However, on these facts, because there were so many separate connections in the chain, the prosecution ought to have obtained evidence from each person responsible for the process by which lost or stolen cards were brought to the attention of the relevant authority, and how the information was disseminated.[1]

1 See 'Evidence obtained from a computer' (1992) 56 Journal of Criminal Law 44–5 for a comparison between *Minors* and *Shephard* and 'touching wood'; Colin Tapper, 'Reform on the Law of Evidence in Relation to the Output from Computers' (1995) 3 Intl J L & Info Tech 85. In *Odex Pte. Ltd. v Pacific Internet Ltd* [2007] SGDC, *rev'd on other grounds*, [2008] SGHC 35, [2008] 3 SLR 18, the lawyers could not even identify the correct person to prepare a witness statement; George Wei, 'Pre-commencement Discovery and the Odex litigation: Copyright versus confidentiality or is it privacy?' (2008) 20 SAcLJ, 591; and Daniel Seng, 'Evidential issues from pre-action discoveries: *Odex Pte Ltd v Pacific Internet Ltd*' (2009) 6 Digital Evidence and Electronic Signature Law Review 25.

Lay experts as witnesses

10.8 In the case of *R v Spiby (John Eric)*,[1] the defence argued, unsuccessfully, that the sub-manager of a hotel could not discharge the burden under s 69 of the Police and Criminal Evidence Act 1984 to show that the computer was working 'properly'. It was submitted that only a service engineer or an expert on the use of the particular computer system would have been able to say whether the machine was working 'correctly'.[2] Taylor LJ agreed with the decision of the trial judge, and considered that the positive evidence of the sub-manager that the device was working was sufficient in this instance. This cannot be correct. Only a service engineer or a suitably qualified professional with knowledge of the particular computer system would be in a position to determine whether the device was working 'properly'. The sub-manager was only competent to give evidence of his reliance on the output of the device for the purpose of submitting a record of the telephone calls made from particular extensions in the

hotel and recorded by the machine – that is, for the purpose of billing customers for the calls made. An assertion that the output is considered reliable because the hotel relies on the output of the device does not prove the device is 'reliable'. These are separate questions.

1 (1990) 91 Cr App R 186, [1991] Crim LR 199, CA.
2 Colin Tapper, 'Evidence from Computers' (1974) 8 Georgia Law Review 562, 595. Professor Tapper noted, at fn 193, 596, that 'An interesting trial dilemma regarding foundation testimony is that too much of a showing of error control may cause a jury to find the system so fraught with error that the system would be presumed to be unreliable, while too little testimony on that matter would cause a similar result.' Unfortunately, it does not follow that the latter result occurs.

10.9 Compare this case with the decision in *United States of America v Linn*.[1] A computer print-out of telephone calls was admitted into evidence. The appellant argued that the print-out was not admissible because it was an untrustworthy record generated by a computer. The appellant suggested that the Director of Communications of the Sheraton hotel '"did not understand the distinctions between 'menus', data bases', and computer 'code', she was 'confused and inadequately trained,'" and thus without personal knowledge of the way in which the computer printout was generated.'[2]

1 880 F.2d 209 (9th Cir. 1989).
2 880 F.2d 209 (9th Cir. 1989), 216.

10.10 No evidence was offered to indicate why the content of the print-out was considered to be unreliable or why it was relevant that the witness failed to understand how the print-out was generated. Beezer CJ rejected the submission as frivolous. He pointed out that the telephone record was generated automatically and it was retained in the ordinary course of business; thus such records were considered business records under the relevant federal rules of evidence.

10.11 In this case, two separate issues were conflated: first, the witness was not an expert witness and therefore not qualified to give the evidence, and second, the witness failed to understand the underlying working of the computer that produced the print-out. If the 'reliability' of the computer was in issue, the appellant ought to have alleged the content of the print-out could not be trusted, and have given sufficient reasons for the burden to fall to the prosecution to demonstrate the computer was working correctly.

10.12 It was not considered necessary for a computer expert to provide evidence that a till roll connected to a computer was 'working properly' in *R v Shephard*[1] under the provisions of s 69 of the Police and Criminal Evidence Act 1984. The oral evidence of a store detective, who demonstrated how the prices of goods were added to the till roll, was considered sufficient by the members of the Court of Appeal and the House of Lords. It should be noted that the store detective was only capable of demonstrating the method by which the prices of goods were added to the till, not whether the software accurately replicated the list of goods purchased. In giving judgment in the Court of Appeal, Lloyd J said of the evidence given by the store detective:[2]

> On the evidence in the court below in the present case, there was no doubt about the functioning of the computer. Mrs. McNicholas who gave detailed evidence as to how the cash tills worked, and explained the link with the central computer, was asked in chief

> 'Q. And what about the master computer? Did that malfunction?
>
> A. Touch wood, no. I have never known it break down since we have had it.'
>
> She was not cross-examined on the point. In addition, she has spent, as we have said, many hours examining the particular till rolls. She would have been the first to notice if there had been any internal evidence of malfunction. In those circumstances it was legitimate for the court to infer that the computer was operating properly.

1 [1993] AC 380, [1993] 1 All ER 225, [1993] Crim LR 295, HL (spelt 'Shepherd' in All ER and Crim LR); but see the highly relevant comments in 'Evidence obtained from a computer' (1992) 56 Journal of Criminal Law 44–5 in comparing the decision in this case against the decision in *R v Minors (Craig); R v Harper (Giselle Gaile)* [1989] 1 WLR 441, [1989] Crim LR 360; 'Admissibility of computer print-outs' (1993) 57 Journal of Criminal Law 277–8.
2 *R v Shephard* (1991) 93 Cr.App.R 139, 143.

10.13 In rejecting the need for a computer expert to sign a certificate where oral evidence has been given that was open to cross-examination, Lord Griffiths offered the following comments in the House of Lords:

> Documents produced by computers are an increasingly common feature of all business and more and more people are becoming familiar with their uses and operation. Computers vary immensely in their complexity and in the operations they perform. The nature of the evidence to discharge the burden of showing that there has been no improper use of the computer and that it was operating properly will inevitably vary from case to case. The evidence must be tailored to suit the needs of the case. I suspect that it will very rarely be necessary to call an expert and that in the vast majority of cases it will be possible to discharge the burden by calling a witness who is familiar with the operation of the computer in the sense of knowing what the computer is required to do and who can say that it is doing it properly.[1]

1 [1993] AC 380, [1993] 1 All ER 225, HL at 387 B–D; followed in *Public Prosecution Service v McGowan* [2008] NICA 13, [2009] NI 1.

10.14 Lord Griffiths went on to say:

> The computer in this case was of the simplest kind printing limited basic information on each till roll. The store detective was able to describe how the tills were operated, what the computer did, that there had been no trouble with the computer and how she had also examined all the till rolls which showed no evidence of malfunction either by the tills or by the central computer.[1]

1 [1993] AC 380, [1993] 1 All ER 225, HL at 387E.

10.15 Dr Stephen Castell was engaged as an expert witness in litigation regarding a major electronic point of sale computer system for a national retailer in 1994, and he remarked that a centralized computer connected to remote tills in store branches is far from being a computer of the simplest kind.[1]

1 'Letter to the Editor' (1994) 10 Computer L & Secur Rep 158.

10.16 At the same time as this case was being heard in England, the Court of Appeals of Nebraska heard an appeal in the case of *State of Nebraska v Ford*.[1] The appellant was convicted of theft from hotel rooms. The hotel used a system controlled by a

computer, by which both those staying at the hotel and members of staff gained entry to a room by way of a card with machine readable code. A number of thefts from rooms were linked to the recorded use of a card issued to Ford. When challenged, Ford admitted to being the rooms at the time, but not to theft. The prosecution adduced the business records under the hearsay exception, which provides that the evidence can be admitted if the activity recorded is of a type that regularly occurs in the course of the day-to-day activity of the business; and the record was made at or near the time of the events recorded, and the record is authenticated by a qualified witness. The defence challenged the qualifications of the witness, Glenda Willmon, the general manager of the hotel, who explained how the system worked. Connolly J, who gave the judgment for the court, rejected the submission by the defence that the witness was not suitably qualified. The judge said that it did not matter whether the witness could discuss the components or engineering principles of the computer.[2] This must be right. Unless there is a challenge to the accuracy of the evidence tendered that results from a computer or computer-like device, it does not necessarily follow that a person familiar with a computer system cannot give evidence of the output of the system.

1 501 N.W.2d 318 (Neb.App. 1993).
2 501 N.W.2d 318, 321.

10.17 The view that an expert is not always required to attest to the proper working of a computer was repeated in *Darby (Yvonne Beatrice) v DPP*.[1] In this case, a police constable operating a speed measuring device testified to the proper operation of the device, even though the device acted to corroborate his own testimony. In undertaking this task, the police constable merely outlined how the device was used, not whether it was accurate. Similarly, in *R v Dean and Bolden*,[2] Lt Cdr Quigley, a Maritime Law Enforcement and Liaison Officer at the Department of State, contacted the Coast Guard Command Centre at US Coast Guard headquarters in Washington, DC to request a search of the vessel 'Battlestar'. A search was made of the Marine Safety Information System, which was a database containing information on all US vessels. The Command Centre also searched the databases of four coast states, and no record of this vessel was found. One ground of appeal centred on the submission that there was no evidence from the people who carried out the searches and the computers were operating properly, and as a result, the evidence was not admissible under s 69 of PACE 1984. The members of the Court of Appeal disagreed. It was considered that Lt Cdr Quigley could give evidence of the 'reliability' of the computers, because there were no reported problems with the databases, and that searches on three separate occasions for the same name failed to bring up the name of the vessel. Dyson J gave judgment, and commented that: 'the fact that searches on three separate occasions produced the same result provided strong support for the conclusion that the computers were operating properly on each occasion'.[3]

1 [1995] RTR 294, (1995) 159 JP 533, DC.
2 (1998) 2 Cr App R 171, CA.
3 (1998) 2 Cr App R 171, 178E.

10.18 This conclusion ought to be reconsidered: the proposition should be that the database was searched on three occasions, and the failure to find an entry for the vessel enables the conclusion to be reached that the name of the vessel was not on the database.[1] This is a different issue as to whether the computer was 'working properly',

or in preference, returning verifiably correct results: the computer may not have been working completely to the expectation of the user, because it might have had any number of problems that did not necessarily affect the effectiveness of the search facility. The effectiveness of the search of the database can be independent of the ability of the computer to return generally verifiably correct results. If the 'reliability' of the computer is challenged, it must be necessary to provide a reasonable basis upon which the claim is made, and there ought to be some evidence proffered to demonstrate the results produced by the computer might be so unreliable as to affect the output used in evidence.

1 *The Queen on the application of Sedgefield Borough Council v Dickinson* [2009] EWHC 2758 (Admin), where a search of a database failed to reveal evidence of an entry, but this was insufficient to prove that the notification of a change of circumstances had not been received.

Qualification of witnesses

10.19 Where there is a reason that the content of the computer print-out cannot be trusted, then the qualifications of the witness will be relevant, because of the nature of the evidence they will be required to give and be cross-examined upon. The degree of expertise required from a witness will vary, according to the problem encountered. In *DPP v Barber*,[1] the first two characters of each line were missing on the print-out. The accuracy of the information recorded on the print-out was not affected. However, the magistrate declined to hear the evidence of a service engineer that was able to explain the nature of the problem because he was not a computer expert and the evidence of what he had seen at a later date was not relevant to the state of the device at the time the print-out was produced. The appeal was allowed because the evidence of the service engineer should have been received. This must be right, given that an ancillary part of the device was apparently not working properly, and the defect did not affect the accuracy of the data.

1 (1999) 163 JP 457; 'Effect of Intoximeter's defects' (1999) 63 Journal of Criminal Law 527–9.

10.20 The two issues are further illustrated in *R v Neville*,[1] where the Crown sought to adduce evidence of a computer print-out showing telephone calls made on Neville's mobile telephone in connection with the hiring of a tractor unit and the employment of a driver to transport a large quantity of stolen hi-fi equipment. The mobile telephone was hired from Talkland, a subsidiary of ICL. A different company, Racal, undertook the telephone operations. The software in the Racal computer issued instructions to record the date, time and duration of each call automatically, and these details were passed on to Talkland. The computer belonging to Talkland included software code that enabled it to produce an itemised bill for their customers. When the bill was paid, the print-out was stored on microfiche. The Crown sought to adduce the microfiche into evidence (or, presumably, a print-out of the contents recorded on the microfiche), and the judge admitted it after a trial within a trial. The Crown then called a witness, an employee of Talkland with no apparent qualifications, to give evidence that she checked all relevant records and had no reason to believe that the telephone bill was inaccurate because of any improper use of either of the computers involved, including the Racal computer. She also stated that the computer at her place of work was working properly so far as her enquiries led. This cannot be correct. The witness might have had the competence

to give evidence of the procedures within her knowledge to provide for the accuracy of billing information at Talkland,[2] but was in no position (not being competent) to offer evidence of any material substance that the computers at Talkland were working properly, and certainly not in a position to offer the same evidence relating to the procedures at Racal, nor as to whether the computer belonging to Racal, of which she had no knowledge, never mind expert knowledge, was working properly.

1 [1991] Crim LR 288.
2 The evidence can be admitted under the provisions of s 117 of the Criminal Justice Act 2003.

10.21 Knowledge that is obtained from experience at work in the absence of formal qualifications is also acceptable,[1] although it is not helpful when a police officer is entrusted to conduct a forensic examination of a mobile telephone without the relevant knowledge or expertise, as in *R v Coultas*.[2] The degree of expertise required of a witness was the subject of the appeal in *R v Stubbs*.[3] The appellant was convicted of conspiracy to defraud, in that he was involved in the fraudulent money transfers from the HSBC Bank of around £11.8m. The fraudulent activities were carried out using an online banking system called 'Hexagon'. The appellant was a member of the password reset team, responsible for resetting customer passwords. The prosecution called Mr Richard Roddy, an employee of HSBC, to give evidence of the Hexagon system. Mr Roddy was not the only witness called to provide evidence of an expert nature. The defence objected at trial to the admissibility of parts of Mr Roddy's evidence on the basis that he lacked the expertise and independence to give expert opinion on the matters in question. It was accepted that he could give evidence about the processes within HSBC and the manner in which the system was designed to operate. However, it was contended that his detailed account of the actual activity within the system at the material times amounted to inadmissible opinion evidence. Following a trial within a trial, the judge ruled Mr Roddy's evidence to be admissible and declined to exclude it under s 78 of the Police and Criminal Evidence Act 1984 or art 6 of the European Convention on Human Rights. The grounds of objection are set out in the judgment of Richards LJ:

> 48. Of particular importance was Mr Roddy's evidence that the activity reports all related to the same session, which had the reference number 'CC000051' and had been registered to the staff delegate identification PWRD on the morning of 24 July 2002. A session number would be allocated upon a user's log-on at a particular terminal. If all the transactions took place within one continuous session and there were legitimate transactions admittedly carried out by the appellant during that session just before and just after the illegitimate transactions, the prosecution could argue with force that the illegitimate transactions must have been carried out from the same terminal; and this also provided strong support for the argument that they must have been carried out by the appellant.
>
> 49. Mr Winter submitted that Mr Roddy did not have the expertise to give such evidence that the activity reports all related to a single session. The fact that they had the same number did not mean that it was a single session. There was evidence from the admitted expert, Mr Danbury, that *concurrent* log-ons (so as to target and hijack a live session) were not possible; but that left open the possibility of *non-concurrent* log-ons to the system under the same session number. This was something that Mr Roddy had not investigated and did not have the technical qualifications to investigate or to answer questions about.
>
> 50. Among the various points made by Mr Winter were these:

> i) The activity reports themselves do not show when log-ons and log-offs occurred. For example, they do not show the undoubted log-off by the appellant at about 17.20. This leaves open the possibility that he had previously logged off at about 17.00, just before the illegitimate activity.
>
> ii) There was no evidence about the appellant's log-on in the morning. Further, although Mr Roddy said that the computer timed out if the session was idle for a period, the evidence was not clear as to how long it needed before a timed log-off occurred. One would have expected a timed log-off when the appellant left the appellant at lunchtime, but there was nothing to show whether there had been a log-off followed by a fresh log-on by the appellant after lunch. In short, there was simply no evidence about when or how the appellant's CC000051 session was created.
>
> iii) Mr Roddy gave evidence that, once a session ended, the next session would not be given the same number again: the number reverted to a pool of numbers available to be allocated by the computer to new sessions. He said in cross-examination that there was a 1 in 100,000 chance of it being reallocated to a different session on the same day. Yet there was evidence of three instances the previous day in which session numbers had been reallocated to other sessions after discontinuance of the session to which they were originally allocated. Mr Roddy was unable to say how this could have happened.
>
> iv) There were other pointers to the illegitimate activity having been carried out by someone other than the appellant. The illegitimate activity involved a random attack on five companies beginning with the letter 'A', whereas the appellant would have known or could have discovered the primary delegate identification for all the companies and would not have needed to do things in this way. Moreover, on two occasions in the course of the illegitimate activity the user deployed a shortcut that was never used by the appellant in the course of his legitimate transactions. The vulnerability of the system to attack by members of staff was illustrated by the fraud perpetrated by Mr Kareer earlier the same year, involving as it did the use of other people's terminals in their absence.[4]

1 *R v Oakley* (1980) 70 Cr App R 7, [1979] RTR 417, [1979] Crim LR 657, CA where a police officer, with 15 years' experience in the traffic division, attended and passed a course as an accident investigator and having attended over 400 fatal road traffic accidents; *R v Murphy* [1980] QB 434, [1980] 2 All ER 325, [1980] 2 WLR 743, CA where a police officer offered an opinion as to the nature of a collision.
2 [2008] EWCA Crim 3261, 2008 WL 5725548.
3 [2006] EWCA Crim 2312.
4 [2006] EWCA Crim 2312, [48]–[50].

10.22 In reaching the decision to admit the evidence, the trial judge applied the tests in *R v Bonython*.[1] Richards LJ agreed that it was not in dispute that the first test was satisfied, because the Hexagon system was a subject for expert testimony, and he went on to say, of the second question:

> In our judgment he was also right to give an affirmative answer to the second question, holding that Mr Roddy had acquired sufficient knowledge of the subject to render his opinion of value in resolving the issues before the court concerning the operation of the Hexagon system. This was an assessment properly made after hearing Mr Roddy's evidence on the *voir dire*. The extent of Mr Roddy's experience of the Hexagon system, as summarised above, enabled him to give valuable assistance on the interpretation of the data taken from the central computer and set out in the activity reports. It was accepted that he was not an IT

> specialist in any wider sense and that his technical knowledge of the system was limited. But this did not preclude his being regarded as an expert to the extent indicated by the judge.[2]

1 [1984] SASR 45.
2 [2006] EWCA Crim 2312, [55].

10.23 The members of the jury were informed of the limitations in the evidence that Mr Roddy was able to give, and it was a matter for them to determine whether they should accept and place weight on his evidence. It was submitted that Mr Roddy's evidence went to admissibility because he was an employee of HSBC and represented the victim of the fraud, and therefore he was not an independent witness. The court rejected this submission. Expertise and independence are separate issues, and it was pointed out that although he made a concession to his lack of objectivity, no attention was given to any feature of his evidence that would support a case of conscious bias or lack of objectivity. Richard LJ indicated:

> In any event it was a matter for the jury to determine whether there was any conscious or unconscious bias or lack of objectivity that might render his evidence unreliable. This was, as the judge said, a matter going to weight rather than admissibility. The circumstances did not warrant a refusal by the judge to admit the relevant parts of Mr Roddy's evidence at all.[1]

1 [2006] EWCA Crim 2312, [59].

10.24 The technical evidence offered by Mr Roddy was not the only evidence of relevance that was led by the prosecution. There was supporting evidence for the prosecution case, for instance: the appellant left the building sometime after 17.00, and returned at 17.27. He claimed he returned to collect his umbrella and that it had been raining, yet the evidence from a CCTV located outside an office a few minutes away from the entrance revealed it was bright and sunny at the material time. The appellant also failed to produce relevant paperwork authorising the change in passwords, lied during his internal interviews, and the evidence he gave to the police when questioned was also inconsistent.

10.25 In addition to the evidence of Mr Roddy, the prosecution also called a Mr Alan Danbury, a computer expert who had been responsible for introducing the system into the UK in the early 1990s, and the manager of the support team until he retired in 2004. During the trial within a trial, the judge also heard evidence from a witness for the defence, a Mr Michael Turner. Mr Turner was not able to provide a report because of a lack of information for a variety of reasons, as set out by Richards LJ:

> … the appellant's workstation had not been retained or imaged; there was no computer running the 2002 version of the Hexagon system which could be analysed; he had been provided with no information as to how the HSBC computers operated or produced the audit logs relied on by Mr Roddy; and he did not have the underlying data from which he could safely reach any conclusion.[1]

1 [2006] EWCA Crim 2312, [44].

10.26 These comments highlight the problems faced by the defence in attempting to elicit co-operation with the victim, when legitimate questions need to be investigated to cross-examine and undermine the evidence of prosecution witnesses. This is a particular problem when challenging a bank, because the defence has a legitimate

interest in challenging the ability of a particular system to withstand an attack or an attempt at subversion. Conversely, the bank cannot, when confronted with evidence that fraud may have taken place, suspend the operation of the system or disrupt it in such a way as to cause it to stop working, no matter how short a time it would take. If a bank were required to pay more attention to the gathering of forensic evidence at a sufficient standard to satisfy criminal proceedings, then they, together with other organizations that may suffer similar attempts, will be either obliged to train employees, or call in suitably qualified experts to conduct an investigation at the time the suspicion is raised. Apart from the added cost and the marginal utility of taking such steps, the victim must decide at the time suspicion was raised whether the integrity of the system will be at issue, which in turn requires the victim to have hindsight of the future challenges.

10.27 In this case, a balance had to be struck between adducing evidence of the system and how it operated within the knowledge of the person responsible for the system at the bank, and whether it was necessary to require a more in-depth analysis from a person expert in the relevant system itself. The dividing line between the need for an expert in the operation of the computer system to give evidence, and the evidence of someone who is familiar with the day-to-day operation of the system is a fine one, and it will depend on the nature of the case as to whether one expert is to be preferred over another.[1] In many cases, as this particular prosecution illustrates, the expert evidence, both internal and external, will not be conclusive. The members of the jury can be appraised of the conflicting technical evidence, and will then be required to consider the technical evidence against the other evidence in reaching their decision. In this instance, it can be argued that the technical evidence, which was not conclusive, was supported by the inconsistencies in the appellant's behaviour.

1 In *RTA v McNaughton* [2006] NSWSC 115, a witness was not permitted or sufficiently expert to give evidence of the position a vehicle was in at the material time.

10.28 Arguably, there is a distinction between the competence, knowledge and qualifications of a witness tendered to give evidence of the trustworthiness of evidence in digital data. If the defence challenges the accuracy of the evidence, it will be necessary to call a witness with relevant competence, knowledge and suitable qualifications to give evidence. The decision in the case of *R v Shephard* must be right, but not because of the rationale offered by the members of the House of Lords. The defence did not challenge the truth of claims made by the witness, only the qualifications of the witness to testify. From the law reports, it appeared that the witness had sufficient knowledge to offer the evidence they did. Had the defence challenged the system to which the till roll was connected, and questioned whether the entire system was trustworthy, including what, if any, errors had been found in operating it across a number of shops connected to a central server, then the witness would not have been competent or qualified to give evidence.

Appendix 1

Guidelines for the search and seizure of evidence in digital form

Australia

HB 171-2003: Guidelines for the management of IT evidence

European Union

Best Practice Manual for the Forensic Examination of Digital Technology (ENFSI-BPM-FIT-01 Version 01 November 2015, European Network of Forensic Science Institutes, Forensic Information Technology Working Group)

Electronic evidence – a basic guide for First Responders Good practice material for CERT first responders (European Union Agency for Network and Information Security, 2014)

Guidelines on Digital Forensic Procedures (European Anti-Fraud Office (OLAF), 15 February 2016)

International Organization for Standardization and the International Electrotechnical Commission [Heading]

ISO/IEC 27037:2012 — Information technology — Security techniques — Guidelines for identification, collection, acquisition, and preservation of digital evidence

ISO/IEC 27041:2015 — Information technology — Security techniques — Guidance on assuring suitability and adequacy of incident investigative methods

ISO/IEC 27042:2015 — Information technology — Security techniques — Guidelines for the analysis and interpretation of digital evidence

ISO/IEC 27043:2015 — Information technology — Security techniques — Incident investigation principles and processes

ISO/IEC 27050 — Information technology — Security techniques — Electronic discovery (DRAFT)

Internet Engineering Task Force [Heading]

IETF RFC 3227 Guidelines for Evidence Collection and Archiving, February 2002

Scientific Working Group for Digital Evidence (SWGDE)

Best Practices for Chip-Off Version: 1.0 (February 8, 2016)

Best Practices for Collection of Damaged Mobile Devices Version: 1.1 (February 8, 2016)

Best Practices for Computer Forensics Version: 3.1 (September 05, 2014)

Best Practices for Digital Audio Authentication Version: 1 (June 23, 2016)

Best Practices for Examining Magnetic Card Readers Version: 2.0 (September 29, 2015)

Best Practices for Examining Mobile Phones Using JTAG Version: 1.0 (September 29, 2015)

Best Practices for Forensic Audio Version 2.1 (June 30, 2015)

Best Practices for Handling Damaged Hard Drives Version: 1.0 (September 05, 2014)

Best Practices for Mobile Phone Forensics Version: 2.0 (February 11, 2013)

Best Practices for Portable GPS Device Examinations Version: 1.1 (September 12, 2012)

Capture of Live Systems Version: 2.0 (September 05, 2014)

Establishing Confidence in Digital Forensic Results by Error Mitigation Analysis Version: 1.5 (February 05, 2015)

Focused Collection and Examination of Digital Evidence Version: 1.0 (September 05, 2014)

Image Processing Guidelines Version: 1.0 (February 8, 2016)

Linux Tech Notes Version: 1.0 (February 8, 2016)

Mac OS X Tech Notes Version: 1.3 (September 29, 2015)

Recommended Guidelines for Validation Testing Version: 2.0 (September 5, 2014)

United Kingdom

ACPO Good Practice Guide for Digital Evidence (Association of Chief Police Officers, Version 5 October 2011; published 2012).

United States of America

Quality Standards for Digital Forensics (Council of he Inspectors General on Integrity and Efficiency, November 2012)

Crime Scene Investigation: A Guide for Law Enforcement (The National Centre for Forensic Science, September 2013)

Searching and Seizing Computers and Obtaining Electronic Evidence in Criminal Investigations (Computer Crime and Intellectual Property Section, Criminal Division, United States Department of Justice, no date)

Electronic Crime Scene Investigation: A Guide for First Responders (US Department of Justice, Office of Justice Programs, National Institute of Justice, 2nd edn, April 2008)

Forensic Examination of Digital Evidence: A Guide for Law Enforcement Special Report US Department of Justice, Office of Justice Programs, National Institute of Justice, April 2004)

Digital Evidence in the Courtroom: A Guide for Law Enforcement and Prosecutors (US Department of Justice, Office of Justice Programs, National Institute of Justice, January 2007)

Best Practices for Seizing Electronic Evidence v.3: A Pocket Guide for First Responders (US Department of Homeland Security, not dated)

Guide to Integrating Forensic Techniques into Incident Response (National Institute of Standards and Technology, Special Publication 800-86, 2006)

Computer Forensics Tool Testing Project Handbook (National Institute of Standards and Technology, Computer Forensics Tool Testing Program, Office of Law Enforcement Standards, 8 June 2015)

Appendix 2

Draft Convention on Electronic Evidence

First published as a supplement to the 2016 issue of the Digital Evidence and Electronic Signature Law Review.

Summary

The Draft Convention is the first treaty dealing with the status of electronic evidence, covering civil and criminal proceedings; the investigation and examination of electronic evidence, and general provisions regarding the recognition and admissibility of electronic evidence from foreign jurisdictions.

Convention on Electronic Evidence

London,

Preamble

[The States signatory hereto],

Considering that the aim of the Drafting Committee is to encourage judges and lawyers to appreciate the concept of evidence in electronic form;

Recognising the value of promoting international co-operation with [the other States that are Parties] to this Convention;

Convinced of the need to pursue, as a matter of priority, a common policy on electronic evidence;

Conscious that the profound changes brought about by the machine and software code (collectively 'digital systems') have altered the means by which evidence is authenticated, in that the medium and the content are no longer bound together as with paper, and that the rules established for paper do not always apply to evidence in electronic form;

Concerned by the risk that electronic evidence can be misunderstood and misinterpreted;

Recognising that evidence in electronic form has unique characteristics that are significantly different to paper and other objects, which raise complex questions about the integrity and reliability of data in electronic form;

Recognising the need to facilitate the co-operation between States for the proper receipt, handling and authentication of electronic evidence;

Believing that it is in the interests of justice to provide for fairness in legal proceedings;

Have agreed as follows:

Part I – Use of terms

Article 1 – Definitions

For the purposes of this Convention:

"adjudicator" means any person that is lawfully appointed as a judge, arbitrator or to any other role that requires the holder of the office to act in a judicious and unbiased manner;

"attribution" means the assigning of responsibility for or tracing the origin of an act purported to have been performed or committed using or through a computer device,

system or network;

"authentication" means the process by which any electronic record, document, statement or other thing is proven to be what it claims to be;

"computer" means any device capable of performing mathematical or logical instructions;

"court" means any international court, national court, statutory arbitral or other tribunal, board or commission according to national law of the contracting state;

"electronic evidence" means evidence derived from data contained in or produced by any device the functioning of which depends on a software program or from data stored on or communicated over a computer system or network;

"electronic record" means data that is recorded or stored on any medium in or by a device programmed by software code and that can be read or perceived by a person or any such device, and includes a display, printout or other output that represents the data;

"device" means any apparatus or tool operating alone or connected to other apparatus or tools, that processes information or data in electronic form;

"digital" means anything that relies on technology based on a binary system or any future development or replacement technology of the same;

"digital evidence practitioner" means a person who is appropriately qualified, and where the law requires, authorized, to investigate and examine evidence in electronic form;

"legal proceeding" means any formal procedure that takes place before any court, national or international, a statutory arbitral or other tribunal, board or commission according to national law and charged with legally defined duties and obligations, or any other formal legal process;

"metadata" means data that describe other data;

"program" means any set of instructions stored in a machine-readable format that can be used to perform a function in a repeatable and reproducible manner;

"relevant legal proceedings" means the legal proceedings for which data in electronic form is requested under a Mutual Legal Assistance Treaty or any other bilateral or multilateral instrument;

"tool" means any device or software program that can be used to identify, secure, examine and analyse electronic evidence.

Part II – Status of electronic evidence

Article 2 – Admissibility of electronic evidence

1. Evidence in electronic form shall be admitted into legal proceedings.

2. Article 2(1) does not modify any existing national rule that applies to the admissibility of evidence, except in relation to the rules relating to authenticity and best evidence.

Article 3 – Agreement on the admissibility of electronic evidence

1. Unless otherwise provided in any law operating in the relevant jurisdiction, an electronic record or document may be tendered, subject to the discretion and rules of the court, if the Parties to the proceedings have expressly agreed to its introduction.

2. Notwithstanding the provisions of Article 3(1), an agreement between the Parties on the admissibility of an electronic record or document does not render the record admissible in a criminal proceeding if at the time the agreement was made

(a) the accused person or any of the persons accused in the proceeding was not represented by a lawyer;

(b) except where the adjudicator finds that admitting the record or document into evidence does not prejudice the case for the accused.

Article 4 – Authentication of electronic evidence

1. The party seeking to introduce electronic evidence in any legal proceeding has the burden of proving it is what it claims to be.

2. The matters set out below are to be considered when assessing that evidence in electronic form is what it claims to be:

(a) The data (both the content and associated metadata) relied upon in any legal proceedings can be shown to be an accurate representation of the prevailing and existing state of those data at the time relevant to the legal proceedings.

(b) If the data have changed from the moment they were identified (and possibly seized) as potential evidence in legal proceedings, there is an accurate and reliable method of documenting any such changes, including the reasons for any such modifications.

(c) The continuity of the data between the moment in time the data were obtained for legal purposes and their submission as an exhibit in legal proceedings can be demonstrated.

(d) Any techniques that were used to obtain, secure and process the data can be tested and shown to have been appropriate for the purpose for which they were applied.

(e) The technical and organizational evidence demonstrates that the integrity of the data is trustworthy, and can therefore be considered reliable and complete (insofar as the data can be complete), which in turn will depend on the circumstances surrounding the data at the time they were identified as being potentially relevant in legal proceedings.

Article 5 – Best evidence

1. In any legal proceeding, where any printout, document or other physical manifestation of the result or output or appearance of any electronic process, record or any other representation of that process or record has been manifestly or consistently acted on, relied upon, or used as the record of the information represented by or stored on the printout, the printout or other physical manifestation shall be considered the best evidence and admitted as evidence subject to satisfactory proof of its integrity.

2. Where the output of a process is relied upon, and it remains in electronic form, the best evidence rule remains, subject to the provisions of Article 4(2).

3. Article 5(1) and (2) do not modify any domestic rule that applies to the admission of evidence.

Part III – Investigation and examination of digital evidence

Article 6 – Digital evidence practitioner

1. Since digital evidence practitioners are required to make informed judgements about the appropriateness of the tools and techniques they use to secure and preserve electronic evidence, the Parties shall establish minimum standards for their formal education and training.

2. A digital evidence practitioner must be able to provide, in compliance with the necessary court and legal requirements:

(a) an analysis of their findings, setting out the scientifically agreed basis upon which their judgement is based; and

(b) shall identify and explain any data that appear to be inconsistent with their findings.

3. The primary duty of the digital evidence practitioner is to the court.

Article 7 – The use of good practice guidelines for electronic evidence

1. The Parties to the Convention shall establish a Forum for the development of good practice and guidelines in the acquisition, handling and otherwise processing of electronic evidence in the form of a set of agreed common requirements.

2. The forum shall:

(a) Include participation from at least two thirds of all Parties to the Convention.

(b) Establish its own rules of procedure and may establish subcommittees to consider specific issues.

(c) Be funded on a basis to be agreed.

(d) Submit the first edition of its agreed common requirements to the Parties within two (2) years of this Convention coming into force for subsequent adoption by the Parties.

(e) Produce updates and amendments to the agreed common requirements as deemed desirable and necessary by the Forum and in any case every two years, or a statement that an update is not currently necessary.

3. Except where incompatible or inconsistent with national legislation, codes or procedure, the Parties to this Convention shall implement agreed common requirements on the acquisition, obtaining, packaging, processing and examination of electronic evidence.

4. The agreed common requirements shall be:

(a) Drafted by reference to the guidelines established by the Forum.

(b) Adopted within [*time period to be agreed*] of accession to this Convention or within [*time period to be agreed*] of the publication of the first version of the agreed common requirements by the Forum, wherever is the sooner.

(c) Implemented by all national and government departments charged with legal duties and obligations involving the use, handling or processing of electronic evidence.

5. Any authority responsible for investigating a matter involving the criminal law shall apply and follow the agreed common requirements unless there are exceptional or extenuating circumstances where they cannot be followed.

6. Where, under Article 7(5) above, the agreed common requirements have not been complied with for exceptional circumstances, those circumstances and the reasons shall be recorded in writing at the time of the departure from the agreed common requirements and the written record shall be admissible in legal proceedings.

Part IV – Treatment of electronic evidence upon receipt

Article 8 – The requesting party

1. The provisions of this Article apply where the requesting party makes a request for evidence in electronic form to the sending party.

2. When the requesting party makes a request for evidence in electronic form, regardless of the mechanism by which the evidence is requested, the requesting party

shall provide a legally binding undertaking in writing to the sending party to include the following:

> (a) An assurance that the data shall be dealt with in accordance with how evidence in legal proceedings is normally dealt in the requesting parties' jurisdiction under the relevant legislation, procedural rules and rules of professional conduct.
>
> (b) Copies of the data shall only be given to parties authorized to receive the data that are part of the relevant legal proceedings.
>
> (c) Data provided under the provisions of this Article 8 shall only be used for purposes related to the relevant legal proceedings.
>
> (d) The sending party may waive the provisions of Article 8(2)(b). The terms of any such waiver shall be decided by the parties in a form and to the extent that they determine.

3. Notwithstanding the provisions contained in Article 8(2) above, all data in electronic form that is provided to the requesting party shall be the subject of all the relevant laws of the requesting party, including, but not limited to, confidentiality, the protection of data and the security of data.

4. The assurances provided by the receiving party under the provisions of Article 8(2) above may be provided in physical or electronic form as is agreed between the parties.

5. The provisions of Article 8(3) shall also apply to any other receiving party authorised to receive the data that are part of the relevant legal proceedings.

Part V – General provisions

Article 9 – Admissibility of electronic evidence from other jurisdictions

1. Where electronic evidence originates in another jurisdiction, its admissibility is not impaired if the electronic evidence is proven in accordance with Article 3 or the authenticity of the evidence is otherwise demonstrated.

2. The provisions of this Article 9 do not modify any domestic rule that applies to evidence in electronic form obtained contrary to relevant human rights legislation or data protection legislation.

Article 10 – Recognition of foreign electronic evidence and signatures

1. In determining whether or not, or to what extent, data in electronic form are legally effective, no regard shall be had to the geographical location where the data were created or used or to the place of business of their creation, provided those data are located in the domestic jurisdiction.

2. Where the electronic record or document is located in a foreign jurisdiction, Article 10(1) above does not apply unless –

> (a) the party who adduces evidence of the contents of an electronic record or document has, not less than 14 days before the day on which the evidence is adduced, served on each other party a copy of the electronic record or document proposed to be tendered, except where exceptional, urgent and exigent circumstances apply;
>
> (b) the court directs that it is to apply; or
>
> (c) there is an international treaty in effect establishing recognition of electronic records or documents or of electronic signatures located in the foreign jurisdiction.

4. Notwithstanding the provisions of Article 10(2)(a) above, what constitutes exceptional, urgent or exigent circumstances for the purposes of this Article is a matter

for the court seized with the matter.

4. Notwithstanding the provisions of Article 10(2) above, an adjudicator may admit data in electronic form that are located in a foreign jurisdiction if domestic law so provides.

Article 11 – Interpretation

1. Where the meaning of a word or phrase in this Convention differs from the meaning of a word or phrase defined in any information technology literature, the adjudicator shall interpret the meaning in accordance with the domestic law on the interpretation of words and phrases.

Article 12 – Entering into force

1. The Convention shall enter into force on the thirtieth day following the date of deposit with the [*name of sponsoring organization*].

2. For each State ratifying or acceding to the Convention after the deposit of the [*third*] instrument of ratification or accession, the Convention shall enter into force on the thirtieth day after the deposit by such State of its instrument of ratification or accession.

Explanatory notes to the Draft Convention on Electronic Evidence

1. The main objective is to pursue a common policy towards electronic evidence, taking into account the differences in the treatment of evidence in individual jurisdictions. This Convention does not seek to harmonize judicial systems. The aim is to encourage judges and lawyers to more fully understand the concept of electronic evidence in the interests of providing for fairness in legal proceedings; to promote adequate procedures in legal proceedings; to implement appropriate legislation where necessary, and to promote international co-operation.

2. Part I Article 1 provides a number of definitions. The aim is to provide definitions that transcend legal cultures. Although the definition of "authentication" does not include reference to relevant international or domestic guidelines and standards, it does not preclude the use of such guidelines and standards in demonstrating authenticity. The definition of "electronic evidence" is taken to be synonymous with the term "digital evidence".

3. Part II considers the status of electronic evidence, covering the admissibility of electronic evidence (Articles 2 and Article 3), authentication (article 4) and best evidence (Article 5).

4. Article 2 aims to provide minimum rules to the admissibility of electronic evidence. The purpose of Article 2(1) is to prevent a party from seeking to exclude evidence in electronic form because it is in electronic form. Article 2(2) does not modify any domestic rule relating to the admissibility of electronic evidence other than in relation to authenticity and best evidence.

5. Article 3, regarding the agreement on admissibility of electronic evidence, is taken and adapted from the *Commonwealth Draft Model Law on Electronic Evidence* and *Electronic Evidence: Model Policy Guidelines & Legislative Texts* (Harmonization of ICT Policies, Legislation and Regulatory Procedures in the Caribbean, International Telecommunication Union Telecommunication Development Bureau, Geneva, 2013).

6. The provisions of Article 3(1) aim to permit the parties to a legal proceeding to

agree on the authenticity of the evidence. The purpose of Article is to simplify the legal process by reducing the time that might be spent in authenticating documents and records in electronic form that both parties rely on. There is no point in increasing the time (and costs) spent on unnecessary actions.

7. Article 4(1), deals with the process of proving that data in electronic form is what it claims to be. The word authenticity is used, even though this may be considered to be irrelevant and out-of-date. To establish whether a electronic record, document or other thing is proven to be what it claims to be, the tests regarding the integrity, reliability and completeness of the data and therefore trustworthiness is more important. It is for the adjudicator to assess the evidence before them to determine whether the data is what it claims to be. The term 'authentic' is used by many jurisdictions in other contexts, such as the provision of an 'authentic' record. The word 'authentication' remains, but it should not be taken to override the domestic methods of determining whether a electronic record, document or other thing is proven to be what it claims to be – nor does it refer to the 'authentic' record.

8. Article 4(2) was initially taken from Stephen Mason, *Electronic Evidence* (3rd edn, LexisNexis Butterworths, 2012), 4.21. Both the *Commonwealth Draft Model Law on Electronic Evidence* and *Electronic Evidence: Model Policy Guidelines & Legislative Texts* (Harmonization of ICT Policies, Legislation and Regulatory Procedures in the Caribbean, International Telecommunication Union Telecommunication Development Bureau, Geneva, 2013) provide for a presumption (the term 'judicial notice' is also used in some jurisdictions – this term has a similar effect to the presumption) that electronic evidence is 'reliable' or that a computer system or other similar device was 'operating properly'. No lawyer or judicial authority has put any evidence forward to establish what 'reliability' means in relation to computers and computer like devices, or what 'operating properly' means. Because a minority of jurisdictions adopts this presumption in the absence of any evidence that such a presumption is justified, it is considered more appropriate to refrain from including such a presumption in the Draft Convention.

9. The provisions of Article 4(2) operate to require a party to demonstrate whether the data in electronic form it is what it claims to be, and conversely, for the challenging party to cross examine to establish that the data is not an accurate presentation of what it claims to be.

10. Article 5 specifically refers to the common law concept of best evidence. The term 'original' has deliberately not been included in this Draft Convention. This is because the word 'original' has different meanings for lawyers and notaries, and also in different jurisdictions. The term 'original' is not helpful when analysing evidence in electronic form. This is because every item of data in electronic form is a copy. There can be no original.

11. Part III deals with the investigation and examination of electronic evidence in Articles 6 and 7.

12. Article 6 provides for the formal education and training of digital evidence practitioners. People that investigate, seize and analyse evidence in electronic form ought to be educated and trained through a formal process. This is in the interests of justice and fairness between the parties, and because evidence in electronic form is now ubiquitous and an every-day part of legal proceedings.

13. Article 7 provides for the creation of a Forum to develop appropriate guidelines

or standards for the process of investigating evidence in electronic form. A number of guidelines exist at present. It is the interests of justice that such guidelines are not only publicly available, but are developed by representatives from internationally respected bodies. By developing a set of internationally recognized guidelines, adjudicators will be better informed when assessing evidence in electronic form. The development of common guidelines or standards will also promote confidence in and acceptance of the quality of evidence especially where obtained in another jurisdiction.

14. Part IV provides for the transmission of data in electronic form between jurisdictions. The terms of Article 8 do not affect the provision of any Mutual Legal Assistance Treaty, bilateral or multilateral instrument, or of any other method of requesting evidence from a foreign jurisdiction. The purpose of this provision is to reassure the sending party that the evidence sent will be dealt with appropriately and in accordance with the norms of the receiving jurisdiction relating to evidence in legal proceedings. Some jurisdictions are wary of sending evidence without suitable provision for the security and the protection of the people mentioned in the data.

15. Part V deals with general provisions. In particular, Article 9 on the admissibility of electronic evidence from other jurisdictions attempts to deal with the difficult question of which set of legal requirements apply to evidence in electronic form – whether it is of the State in which the evidence is geographically located, or the State in which the evidence is to be submitted in a legal proceeding. Article 9(1) seeks to indicate that if the evidence is proven in accordance with the provisions of Article 4, the matter of the geographical location is irrelevant. Alternatively, an adjudicator can admit the evidence as being authentic where the authenticity of the evidence is demonstrated in some other manner that is accepted by the adjudicator.

16. Article 10 provides that evidence in electronic form that ostensibly originates in a foreign jurisdiction can be admitted, notwithstanding that it was not actually located in the domestic jurisdiction. The aim is to enable the admission into a legal proceeding of electronic evidence and electronic signatures that might otherwise not be admitted because of lack of formalities.

17. Although the provisions of Article 11(1) may appear to be open to interpretation, the clause mirrors many such clauses in legislation relating to electronic commerce and communications across the world. Article 11(2) deals with the inevitable disagreement between the meaning of words in a technical sense and a legal sense. When this occurs, it is for the adjudicator to determine the meaning in accordance with the relevant provisions in domestic law on interpretation. There has been no attempt to incorporate technical definitions into the Convention, because doing so might cause greater uncertainty than is intended.

For the history of the Draft Convention on Electronic Evidence, the reader is referred to the supplement to the 2016 issue of the *Digital Evidence and Electronic Signature Law Review*. DOI: http://dx.doi.org/10.14296/deeslr.v13i0.2321.

List of participants (online and offline)

Carmelo Asaro, retired Italian judge, teaching courses on the degree of Master sulla Sicurezza and Master sul Cyberceime at Departimento di Informatica in the Università degli Studi di Roma 'La Sapienza', Rome

Steven David Brown, Independent law enforcement consultant

Hein Dries, LLM

Dr Mark Lomas, Capgemini UK plc

Dr Steven J. Murdoch, Royal Society University Research Fellow in the Information Security Research Group of University College London

Dr. iur., associate professor Uldis Ķinis, Rīgas Stradiņa Universitāte

Tim McCormack

Angus M. Marshall, BSc, CEng, FBCS, CITP, FRSA, Director and Principal Scientist, n-gate Limited; Director, Digital Evidence Virtual Centre of Excellence C.I.C. and Visiting Fellow at the Open University

Goran Oparnica, Managing Director of INsig2 d.o.o.

Bertan Özerdağ, Judge of the Kuzey Kıbrıs Türk Cumhuriyeti Yüksek Mahkemesinin (Supreme Court of the Turkish Republic of Northern Cyprus)

Gita Radhakrishna, senior lecturer at the Faculty of Law, Multimedia University, Malaysia

Dr Giuseppe Vaciago, Partner at R&P Legal and Lecturer at the Faculty of Law, University of Insubria (Como), Italy

Events

Launch of the Draft Convention on Electronic Evidence

Held at DataFocus 2016, Zagreb, Croatia, 5 April 2016

Workshop on the Draft Convention on Electronic Evidence

Held on 20 May 2016 between 14:30 and 17:00 at the Institute of Advanced Legal Studies, 17 Russell Square, London WC1B 5DR

Attendees

Michael Asher, Barrister

Werner R. Kranenburg, Attorney and Counselor-at-Law, Krenenburg

Dr Alan McKenna, Associate Lecturer, Law School, University of Kent

Naraindra Maharaj, Datatec Financial Services Limited

Nikolaos Trigkas, LLB, MBA, PhD in Law candidate (University of Aberdeen)

Katrine Broch Petersen

Dr Michael Reynolds, Solicitor and Arbitrator

Dr Judith Townsend, Director, Information Law and Policy Centre, Institute of Advanced Legal Studies, London

Richard Trevorah, tScheme Limited

Acknowledgments

A brief introduction to the development of this Convention can be read here: Stephen Mason, 'A proposed Convention on Electronic Evidence', *Pandora's Box*, 2016, 153 – 155 (http://www.jatl.org/pandoras-box/). I was invited by the L'Accademia di Diritto Europeo – Academy of European Law – Europäische Rechtsakademie – l'Académie de droit europée to speak at an event entitled 'Relying on Electronic Evidence in Criminal Cases' (event number 315DT21) held in Bucharest on 12 and 13 November 2015 at the Institutului National al Magistraturii. One of the attendees asked a question that is often asked at similar events: 'Why was there no Convention on Electronic Evidence?'. My usual response was that no organization wanted to spend the time developing one, but on this occasion, I decided at this event to write one myself, and announced that this is what I was going to do.

Part of the content of this Draft Convention on Electronic Evidence was taken from the *Commonwealth Draft Model Law on Electronic Evidence* and the *Commonwealth Draft Model Law on Electronic Evidence* and *Electronic Evidence: Model Policy Guidelines & Legislative Texts* (Harmonization of ICT Policies, Legislation and Regulatory Procedures in the Caribbean, International Telecommunication Union Telecommunication Development Bureau, Geneva, 2013). These valuable sources are explicitly recognised, as is their copyright. I wrote the remainder of the first version of the text.

I am not technically competent, so I was very fortunate that Hein was able and willing to host the web site using the domain name I registered for the purposes of the development of the Convention (conventiononelectronicevidence.org).

My first thanks go to Hein for taking on this arduous task while continuing to work his way around Europe fulfilling various contracts, and also commenting on the content of the Convention.

I also thank the Institute of Advanced Legal Studies for hosting the workshop held in London. It was a useful event.

A final word of thanks to everyone that took the time to read the various iterations of the Convention and offer comments. As can be imagined, lawyers and technicians tend to use language in different ways, and the discussions partly reflect this. I have approached the task of redrafting text by taking into account these differences, and adjusting words where they can be adjusted to the benefit of the project without loosing meaning.

Some suggestions have been made that do not appear in this draft Convention. Their failure to appear is not because they were irrelevant. In drafting a Convention, it is important to ensure that the text can be generally agreed. This means excluding controversial provisions that are not universally shared.

Copyright of the Notice

Appendix 3

Cumulative vignettes

In each edition I have written a short vignette to illustrate a particular issue. Each of those appearing in the first three edition are set out below.

First edition, 2007

The abacus

'Your honour, I seek to exhibit the abacus.'

The judge looked over his spectacles 'Which form of abacus is it?'

The barrister looked perplexed and turned to his solicitor and whispered 'Which form of abacus? How do I know? Are there different types of abacus?'

'Oh yes' whispered the solicitor, 'it's a Chinese abacus.' 'Oh, right. Thanks.' 'It's a Chinese abacus, your honour.'

'Thank you, Mr Puffington. And what is the purpose of exhibiting the abacus?'

'Well, your honour, it's the item upon which the calculations were made to perpetrate the alleged fraud.'

'Indeed, but that does not mean the abacus ought to be exhibited. Have you a submission on this matter Miss Jawleyford?'

Miss Jawleyford stood as Mr Puffington sat down.

'Well, your honour, the defence does not seek to argue about an inanimate object.'

'Quite.'

'But what we must look to, in my submission, is the reason for admitting the abacus as an exhibit, your honour.'

'Indeed.'

'We have already had the opportunity of viewing the abacus, and take no point on the object itself. It is admitted that the defendant used the device. As a material object, it can be admitted into evidence. But the question is, what purpose is served in admitting the device. It is my submission that the presence of the abacus serves no purpose, because the device is merely a device. There is no record of what, if any, calculations might have been made on the device.'

Miss Jawleyford sat down. Mr Puffington stood.

'Your honour, in our submission, it's important to exhibit the abacus, because it will serve to make the members of the jury ask themselves why the defendant, a finance director earning over a million pounds salary a year, deliberately used such a device. It is our case that he used the abacus to avoid the creation of records that would implicate him in the alleged fraud. To that end, it's an important exhibit that ought to be admitted into evidence.'

Second edition, 2010

The 'forged' document

'The problem with the email submitted by the witness, madam, is that the signature cannot be trusted. For this reason, the evidence cannot be admitted.'

Mr Tulkinghorn sat down. Mr Tangle stood up.

'With the deepest possible respect, madam, my learned friend has let his usual penetrating insight into the analysis of evidence fail him. If this was a letter, for instance, the first question will be "Is the letter genuine?" If the letter is a forgery, then the signature matters not – unless it is genuine and intended to deceive the recipient. If the letter is genuine, *then* the question arises as to whether the signature is a forgery. Thus it must be with the email. If my learned friend claims that the email is a forgery, the status of the signature is irrelevant. Is my learned friend suggesting that the email is a forgery?'

Mr Tangle sat down.

Her Honour Judge Flite QC looked at Mr Tulkinghorn 'Well? It strikes me that this must be correct. Are you suggesting the email is a forgery?'

Mr Tulkinghorn stood up.

'In this instance, my learned friend has indicated an error of logic on my part, which I concede. The point is, anybody can forge an email and write any name as an electronic signature. If we cannot trust the signature, then we cannot trust the email.'

Her Honour Judge Flite QC continued the questioning, 'But the authenticity of the email must come before the verification of the signature? Mr Tangle?'

Mr Tulkinghorn sat down. Mr Tangle stood up.

'Where the authenticity of a document is challenged, a wide range of tests can be made to determine whether it is a forgery. I acknowledge that the contents can help determine whether it is a forgery. But if it was a letter, the paper, ink, and the type face might all be the subject of tests. In the case of an email, the technical information relating to the status of the document is of the utmost relevance. In my submission, determining whether to trust the signature can only follow *after* it has been established whether the email is genuine or a forgery.'

Third edition, 2012

The 'competent' witness

'My learned friend for the prosecution has established that you are the sub-manager of the hotel, that you are familiar with the functions of the machine that controls the telephone system, and that you know how it works and what it is supposed to do?'

'Yes.'

'And the print-outs you have brought to court purport to indicate when the telephone was used in room 2820?'

'Yes.'

'For this reason, my learned friend considers your evidence is all that is needed to establish the reliability of the telephone system. Let me ask you this, how does the direct inward system access work?'

'Er, I don't know.'

'You don't know what happens, or you don't know what the direct inward system access is?'

'I don't know what it is.'

'So, by implication, you don't know what the password is?'

'No.'

'By implication, you won't know if thieves have used the password to route telephone calls through the hotel telephone system?'

'No.'

'Can you tell me the purpose of the latest software up-date, whether it included a security fix, and when it was downloaded?'

'Er, no, I don't know any of that.'

'Why do you not know?'

'Well, because the IT people do all of that stuff.'

'So you are asserting, by bringing along the print-outs of the telephone calls, that these telephone calls were actually made, and they were made from room 2820.'

'Well, yes, if you say so.'

'I do not say so, you do. You also claim that because none of your customers have ever complained about their bills, it follows that the telephone system is reliable and therefore trustworthy?'

'Well, I wouldn't put it quite like that.'

'Thank you, Mr Prunsquallor.'

Judge Sepulchrave turned to prosecuting counsel, 'Unless you have any questions in re-examination Mrs Groan?'

Mrs Groan stood up. 'You honour, no,' and sat down.

'Very well, you may leave the witness stand, Mr Prunsquallor. Yes, Mr Rottcodd?'

'Thank you, your honour. My learned friend for the prosecution would have us believe that because the information printed on the piece of paper apparently looks sensible, it therefore follows that the contents must not only be reliable, but represent the truth. My learned friend also suggests that because Mr Prunsquallor uses the hotel's telephone system in the performance of his duties, this is a sufficient foundation as a qualification as a competent witness. With your honour's leave, I will address the latter point first …'

Index

OBserving Law – IALS Open Book Service for Law

In this updated edition of the well-established practitioner text, Stephen Mason and Daniel Seng have brought together a team of experts in the field to provide an exhaustive treatment of electronic evidence. This fourth edition continues to follow the tradition in English evidence text books by basing the text on the law of England and Wales, with appropriate citations of relevant case law and legislation from other jurisdictions.

Stephen Mason (of the Middle Temple, Barrister) is a leading authority on electronic evidence and electronic signatures, having advised global corporations and governments on these topics. He is also the author of *Electronic Signatures in Law* and editor of *International Electronic Evidence*, founding the innovative international open access journal *Digital Evidence and Electronic Signatures Law Review* in 2004. Stephen is an IALS Associate Research Fellow and Visiting Lecturer at the School of Law, University of Tartu, Estonia.

Daniel Seng (Associate Professor, National University of Singapore) teaches and researches on information technology law, infocommunications law, evidence and procedure, artificial intelligence, machine learning and legal reasoning. His research interests also include empirical legal studies and quantitative research and data analytics on big data sets. Between 2001 and 2003, he was concurrently the Director of Research, Technology Law Development Group at the Singapore Academy of Law. Daniel is also a special consultant to the World Intellectual Property Organization, where he has researched, delivered papers and published monographs on copyright exceptions for academic institutions, music copyright in the Asia Pacific and the liability of Internet intermediaries. He is also a non-residential fellow with the Centre for Legal Informatics (CodeX), Stanford University.

With a team of expert contributors.

This book is also available online at http://ials.sas.ac.uk/digital/humanities-digital-library/observing-law-ials-open-book-service-law.

Cover image: fingerprint from Trifonenkolvan/Shutterstock.com

Humanities Digital Library

IALS INSTITUTE OF ADVANCED LEGAL STUDIES | SCHOOL OF ADVANCED STUDY UNIVERSITY OF LONDON