Answers to MoJ Call for Evidence of 2025-01-21 on the current legal presumption that computers are reliable

Peter Bernard Ladkin, Bev Littlewood, Steven J. Murdoch, Martin Newby, Martyn Thomas

2025-04-10

The questions in the Call for Evidence are reproduced below in bold font style. Our contributions as answers are in regular font style.

Questions:

1) The current common law (rebuttable) presumption is that computers producing evidence were operating correctly at the material time.

(a) Is this presumption fit for purpose in modern criminal prosecutions?

No, it is not.

(i) Please specify why you gave this answer

Reasons are given in detail in Peter Bernard Ladkin, Bev Littlewood, Harold Thimbleby and Martyn Thomas, 'The Law Commission presumption concerning the dependability of computer evidence', 17 *Digital Evidence and Electronic Signature Law Review* (2020) 1-14 https://journals.sas.ac.uk/deeslr/article/view/5143

Summary: the presumption is not consistent with the known science of the reliability of software (referred to by practitioners mostly as "software reliability engineering". We shall henceforth use this term).

(b) How easy or difficult do you believe it is at present for this presumption to be effectively rebutted?

It appears to be quite difficult if disclosure is not made of documentation relevant to assessing the reliability of the computer system providing the evidence in the pertinent instances. This was clearly shown in the trial of Seema Misra, in which the Post Office failed in its duty of disclosure of known pertinent system failures. Disclosure applications by Ms. Misra were rejected multiple times by judges. See Stephen Mason, 'The UK Post Office Horizon IT scandal, Part 2: the legal issues', (2024) 30 *Computer and Telecommunications Law Review*, Issue 4, 96-101.

(c) What barriers do you see in effectively rebutting this presumption?

One common problem is inadequate documentation of computer system reliability in the case being considered. Another is inadequate requirements for documentation pertinent to assessing the reliability in the case being considered.

(i) Please give examples where possible.

All of the Horizon trials, as determined by Bates v Post Office Ltd Rev 1 (Horizon Issues) [2019] EWHC 3408 (QB), available at

https://www.bailii.org/ew/cases/EWHC/QB/2019/3408.html . According to our colleague James Christie (who we understand will also be presenting evidence in response to this Call), the accounting functions of Horizon did not appear to conform with standards published by the accounting profession; neither were any assessment activities performed ("internal audit") adequate to demonstrate conformance with these standards (personal communication). This observation is consistent with evidence presented to the Williams Inquiry by forensic accounting software specialist Charles Cipione. We believe appropriate documentation showing assessment and conformance of Horizon accounting functions to the standards set by professional organisations should have been made available to the (many) courts, but it appears it was not. (Let us emphasise here that such general assessment does not and cannot preclude errors being made by the software. But it does illustrate what care might have been taken to exclude such errors.)

The reliability characteristics of software-based systems are particularly poorly understood in the medical profession, and this is so across different jurisdictions. For a Welsh example, see Harold Thimbleby, Misunderstanding IT: Hospital cybersecurity and IT problems reach the courts, 15 *Digital Evidence and Electronic Signature Law Review* (2018) https://journals.sas.ac.uk/deeslr/article/view/4891 The first author (PBL) has recently aided pro bono in the defence before a German court (Bielefeld) of a doctor accused of a crime by entering incorrect information into a patient's electronic medication record. The prosecution presented no evidence from the computer system in question that this had happened (no archive; no back up; no logs). The record was then transcribed manually, then that transcription was sent or transferred to another hospital, where the information in the third medication record in that second hospital turned out to be incorrect. The lack of evidence that the defendant originally made a mistake did not appear to be recognised or challenged by the court (and PBL does not believe that the court record will clarify this, when it is available. German courts have a means of terminating prosecutions with consent of judge, prosecutor and defence; this is what happened in this case and this is what the record will say).

2) Are you able to provide examples from other jurisdictions or situations where the reliability of software must be certified?:

a) As examples of good practice?

Safety-critical software must be developed with specific assessment evidence and specific documentation (50-60 separate documents are required) according to the international functional safety standard for software IEC 61508-3:2010. The intent of IEC 61508-3 is to assure the reliability of software used to implement safety functions. It is best practice for developers of such software to outsource the assessment of their product. Germany has a number of assessment organisations, generally called TÜV, which will perform such a task. Some of the TÜVs are active internationally. Some countries in the European Union include the standard for industrial-process-plant safety, IEC 61511, explicitly in regulations. IEC 61511 refers specifically to IEC 61508-3 for software reliability. In such countries, for such applications, IEC 61508-3 is thus regulation. (The UK is not such a country; HSE considers a developer to have performed due diligence if development is conformant with

IEC 61508-3; and conversely to have been possibly negligent if not. A prosecution for negligence or gross negligence may follow in such a case. This is currently the UK approach to legal proceedings involving software development and safety-related systems.)

b) As examples of things to be aware of?

No comment.

3) If the position were to be amended, what in your opinion would be the most appropriate and practicable solution given our aims and objectives set out above? It would be helpful if your answer could address as many of the below as possible:

A proposal for the probity of computer evidence was made by a group including all of us except SJM to the Ministry of Justice in 2020 pursuant to a request from Minister Chalk to Barrister Paul Marshall. We propose this as the most appropriate and practicable solution: Paul Marshall, James Christie, Peter Bernard Ladkin, Bev Littlewood, Stephen Mason, Martin Newby, Jonathan Rogers, Harold Thimbleby, Martyn Thomas CBE, 'Recommendations for the probity of computer evidence', 18 *Digital Evidence and Electronic Signature Law Review* (2021) 18-26, https://journals.sas.ac.uk/deeslr/article/view/5240 (We refer to this henceforth as *Marshall et al*.)

a) What procedural safeguards need to be in place to ensure your proposed solution is effective?

All the documentation named by *Marshall et al*. should be disclosed to those involved in any legal proceedings. Absence of any specific documentation may be construed as evidence that the computer system has missing documentation supporting the assessment of its reliability.

The documentation named by *Marshall et al*. should be assessed by expert witnesses in software reliability engineering.

b) How might we ensure that any proposed solution is, as far as is reasonable possible, future-proofed?

We offer no general proposal. However, the elephant in the room is the inclusion of so-called "AI" modules, particularly those based on Large Language Models, because assessing the reliability of software which uses such modules is currently an open question in software reliability engineering.

c) How might we ensure that any proposed solution is operationally practical?

The documentation in *Marshall et al*. is operationally practical as it stands.

d) If your proposed solution requires the use of expert witnesses (either jointly or singly instructed), what expertise and qualifications would that person require? To your knowledge are there sufficient such

people at present?

PhD-level academic qualification in software engineering with experience in software reliability engineering, and/or equivalent industry experience of at least a decade to the level of senior engineer and/or activity as an expert witness in legal proceedings and which testimony has been assessed by other software reliability experts as adequate and technically appropriate.

We know a number of internationally-renowned software reliability experts who to our knowledge have never been asked for their expert opinion in legal proceedings. It follows there is an unused pool of software reliability engineering expertise. We would anticipate that this pool would be adequate, at least at first, to fulfil demands on the legal system made by a modification to the law to require probity assessments according to *Marshall et al*. If demand is greater, professional societies can be asked by the MoJ to provide familiarisation training to appropriately qualified professionals on the assessment tasks required.

4) In your opinion, how should 'computer evidence' for these purposes be best defined?

We consider the definition introduced by Baroness Kidron in a proposed amendment discussed in the House of Lords to be appropriate. She used the term "electronic evidence"; we suggest that this is how "computer evidence" be defined; we also prefer the term "electronic evidence", for the reasons given below the proposed definition

[begin definition]

"computer" means any device capable of performing mathematical or logical instructions;

"device" means any apparatus or tool operating alone or connected to other apparatus or tools, that processes information or data in electronic form;

"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 a computer, device or computer system or communicated over a networked computer system."

[end definition]

The reason we prefer "electronic evidence" is that "computer evidence" could be construed to refer to evidence about computers, such as that they are reliable in the technical sense of this term, or that they were designed to behave in such-and-such a way. This is not the same notion as that of "electronic evidence" that we suggest above.

a) Do you agree that evidence generated by software, as set out above, should be in scope, and that evidence which is merely captured / recorded by a device should be out of scope? Please provide a rationale for your answer

Yes, evidence generated by software should be in scope. If evidence is "merely"

captured/recorded by a device then the reliability of that device in capturing/recording the evidence could be in scope if the device is sufficiently complex. Expert assessment of the complexity and reliability of the device should be undertaken.

i) Can you provide specific examples of the type of evidence you believe should be in scope?

No comment.

ii) Can you provide specific examples of the type of evidence you believe should be out of scope?

No comment.

5) Are there any other factors which you believe are important for us to consider?

No comment.

Notes on the Authors

Peter Bernard Ladkin is Professor i.R. (British equivalent: emeritus) of Computer Networks and Distributed Systems at Bielefeld University in Germany. He is a native British citizen. He is Managing Director, resp. Geschäftsführer of the tech transfer companies Causalis Limited in England & Wales and Causalis Ingenieurgesellschaft mbH in Germany. He is Convenor of the IEC Project (now: Maintenance) Team which devised specific development and assessment guidance for software of the highest reliability category arising in the functional safety standard IEC 61508, namely SC4, which was published by the International Electrotechnical Commission (IEC) as IEC TS 61508-3-2 in August 2024.

Bev Littlewood is Professor Emeritus of Software Engineering at City St. George's, University of London and founding Director of its Centre for Software Reliability. He is the 8th recipient in 2007 of the IEEE Harlan D. Mills Award, established "*to recognize researchers and practitioners who have demonstrated long-standing, sustained, and impactful contributions to software engineering practice and research through the development and application of sound theory*".

Steven J. Murdoch is Professor of Security Engineering and head of the Information Security Research Group of University College London. He works on payment system security, privacy-enhancing technologies, online safety, and the interaction between computer science and the law.

Martin Newby is Professor Emeritus of Statistical Science at City St George's University of London. While at City St George's he also held the chair of Product Reliability at the Technological University Eindhoven in the Netherlands. He has worked extensively with notable industrial partners. His work addressed the management of risk and reliability through the development and application of statistical software through the lifecyle of a product.

Martyn Thomas is Emeritus Professor of IT at Gresham College and has testified as an

expert witness in several large IT litigations in the UK, the EU and Australia. He was awarded a CBE for Services to Software Engineering in 2007.

PBL and BL have worked together for a decade and a half on revising the information on statistical assessment of software included in the functional safety standard IEC 61508 for software-based safety-related systems (specifically, IEC 61508-7: 2010 Annex D).

Since 2019 the authors have, with others, including those named in citations above, published a series of studies of aspects of computer evidence before courts, in particular relating to Horizon.