Probabilistic AI Systems in Law Enforcement Futures

PROBabLE Futures Submission in response to Call for Evidence on the 'Use of Evidence Generated by Software in Criminal Proceedings'

Dr. Temitope Lawal, Dr. Angela Paul, Dr. Kyriakos N. Kotsoglou, Professor Marion Oswald MBE (Northumbria University Law School) and Professor Carole McCartney (University of Leicester Law School)

14 April 2025

This is a submission in response to the Ministry of Justice's Call for Evidence on the 'Use of Evidence Generated by Software in Criminal Proceedings' published on 21 January. The submission has been prepared by a team of academic researchers with extensive expertise in the ethical, legal, safeguarding, and operational applications of data analytics and AI within the criminal justice system. Our work is anchored in the UKRI Responsible AI (RAi) UK Keystone Project, "PROBabLE Futures", a four-year research initiative focused on evaluating probabilistic AI systems across the criminal justice sector.

Summary of Recommendations

- **Taking 'Computer Evidence' Seriously:** The original provision from the *Police and Criminal Evidence Act 1984* (PACE) s.69 as well as the succeeding presumption introduced by the Law Commission are simply outdated for the technological advances of today. Many aspects of the criminal justice system are implementing probabilistic algorithmic models (AI systems that uses machine learning to make decisions based on probabilities or likelihoods), with the uncertain outputs being used in evidentiary proceedings. This goes beyond the computerisation of records and production of documents. The various uses of software-generated output include biometric technologies to identify suspects, Generative AI to produce witness statements and crime reports, and risk-prediction tools which can be used to identify individuals and areas of interest to the police.

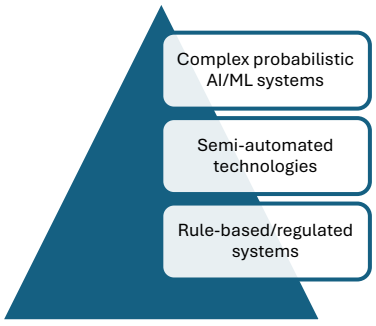

Figure 1: Evidential hierarchy

We therefore propose that the terms 'digital evidence' and 'software-generated digital evidence' be used instead of 'computer evidence'. Here, software-generated digital evidence refers to digital evidence that is the direct output of a computer process (as opposed to digital evidence that is a recording of human action or natural

phenomena). We suggest that the statute or guidance could define "software-generated digital evidence" as *"information intended to be relied upon as evidence of a fact, which information was produced by the application of computer software or an automated algorithm, rather than by direct human perception or input."* Specifically, we propose a **three-tier hierarchical framework** (See *Figure 1* above) – complex probabilistic systems, semi-automated technologies, and regulated, rule-based devices – designed to calibrate legal scrutiny to the complexity, transparency, and regulatory oversight of each technology.

- **Shifting the Burden of Proof to Prosecution and Expert Witnesses:** The current presumption undermines the presumption of innocence, as illustrated in the Post-Office's Horizon technology scandal. Currently, the respective computer system is assumed to be right, unless the defendant can prove otherwise. Many algorithmic systems operate as a 'black box'[1], and machines can be constantly learning. Even a software expert may find it challenging to explain the workings of an algorithm, let alone a defendant who may lack such expertise, thereby posing challenges of evidentiary reliability. Whilst reverse burden of proof clauses (also known as reverse onus clauses) are not per se unlawful, this particular one is not only unmanageable for the defendant but deleterious for the legal system. This became clear recently from the inability of the sub-postmasters to rebut the presumption (that Fujitsu's Horizon system was working correctly) and prove their innocence. It is thus important to shift the burden of proof on software-generated digital evidence. This would also involve emphasising the need for prosecution to acknowledge that errors of probabilistic technology constitute important exculpatory evidence. With the proven biases and errors associated with software-generated outputs, including biometric recognition technologies, it is important that the focus of proving the reliability of a system should be redirected to software developers and to the prosecution. This includes the need for the Ministry of Justice to ensure specialist expertise in software engineering, digital forensics, and related fields for court cases involving software-generated outputs.

- **Increased Scrutiny of Digital Evidence:** There should be robust disclosure obligations regarding the reliability of digital systems used to produce evidentiary outputs. The manufacturer of the technology, the user (the police or the prosecution) should provide reports of errors or problems, maintenance and update records, transaction logs, and information on whether the system has any remote access or human interventions. This increased scrutiny should also extend to courts, incorporating pre-trial case management processes for the assessment of evidence, such as reliability hearings where the prosecution must produce evidence from software-generated outputs. As a result, the trial process involving digital evidence can be streamlined.

- **Auditability and Record-keeping:** To ensure that the evidential output generated by the program is functioning as required, the system (specifically the complex probabilistic ones) should undergo a quality assurance or certification process. A

[1] 'Black box' refers to opacity associated with how an algorithmic system works. For example, when the input and output of an AI system can be seen, but not how the system generated the output from the input.

detailed record of logs, audit trails, and accuracy tests can be presented in court to ensure accountability for how the system operates and produces outputs. It is important that this audit is conducted independently. An example of ensuring this legitimacy could be through certifications from industry standards. Therefore, there is an urgent need for an overarching regulatory framework (either through the establishment of a new regulatory body or the restructuring of existing bodies) to provide central oversight, accreditation, and continuous monitoring of algorithmic tools used in evidentiary contexts. The level of scrutiny on digital evidence should be dependent on the complexities and ambiguities associated with the technology. In addition, if documents have been produced using an algorithmic system, such as a police incident report or a witness statement, there should be an obligation to include a statement that clarifies the use of the algorithmic system (such as a Generative AI tool).

- **Monitoring and Continuous Improvement:** It is important that the findings from the reforms to digital evidence are continuously monitored and assessed. Adequate reviews ensure that our legal system will not fall behind the advancements of emerging technologies. As a result, reforms should not be stagnant. As we have seen, the Law Commission's presumption remained in place for many years and led to miscarriages of justice. The implementation of reforms entails significant resource implications. The expertise required to facilitate adequate knowledge of new technologies, including judicial training, can be costly. Therefore, the reforms should be well-planned in terms of budgeting. It is not enough to implement a reform; those who are affected by the reforms should have the adequate resources in order for the reform to be effective.

1. Introduction

The Post Office Horizon scandal – in which hundreds of sub-postmasters were wrongly prosecuted based on a flawed IT system – illustrates the dangers of presuming digital evidence to be infallible. It is now widely recognised that the current legal presumption that "the computer is always right" (unless proven otherwise) has contributed to serious miscarriages of justice. This submission, focusing on Questions 1, 2, 3, 4, and 5, critiques that presumption's weaknesses, recommends reforms to safeguard fairness and transparency, addresses the need for futureproofing (especially with emerging AI-driven evidence), and advocates adopting modern terminology ('digital evidence' rather than 'computer evidence') in line with technical and jurisprudential developments. Each section below is numbered and headed for clarity, with hyperlinked references to legal and technical sources.

2. The Current Presumption of Reliability and its Weaknesses (Question 1) and (Question 2)

2.1. Presumption of correct operation – origin and rationale

In England and Wales under the common law, it is a presumed fact that a computer system generating evidence was operating correctly at the material time, absent evidence to the

contrary. In practical terms, any output produced by a computer or software is treated by default as reliable unless the defence challenge the system's functioning. This rebuttable presumed fact, introduced by the Law Commission, replaced a stricter regime under PACE s.69, which dealt with 'evidence from computer records'. Under the previous regime, a "statement in a document produced by a computer shall not be admissible as evidence of any fact", unless it could be shown that there were no reasonable grounds for believing that the statement was inaccurate due to improper use of the computer; that the computer was functioning properly at all material times; and that if the computer was not operating properly, it would not affect the production of the document.

Although it was a stricter regime, the original provision was a product of its time. For instance, the title of the section, "evidence from computer records" likely reflects the movement toward the computerisation of records and documents in the 1980s. S.69 PACE was repealed in 1999 following the Law Commission's recommendation, based on the assumption that a general common law presumption of proper functioning would suffice. The Law Commission optimistically believed that such a presumption "would work fairly", assuming that only minimal evidence would be needed for rebuttal, and that courts would not convict if a defendant lacked the means to challenge a system. This approach drew on an older principle that mechanical instruments (like intoximeters) should be presumed "in order" unless shown otherwise. However, treating complex software systems as if they were simple mechanical devices has proven profoundly problematic. This is not an accurate representation of the outputs that can be used in courts nowadays, including probabilistic outputs produced by machine learning algorithms. Thus, there should be a shift away from using ambiguous and outdated terms such as "computer evidence" (as we discuss later in this submission).

2.2. Reversing the burden of proof – what can we learn from empirical studies and caselaw?

The chief concern is that the aforementioned presumption effectively reverses the burden of proof, undermining the presumption of innocence. Normally the prosecution must prove to the requisite standard that evidence is accurate and reliable. Under the current rule, by contrast, courts assume accuracy and the onus shifts to the accused to *disprove* the reliability of the computer evidence. As Mason observes, *"It says, for the person who's saying 'there's something wrong with this computer', that they have to prove it. Even if it's the person accusing them who has the information".* This is a clear departure from the golden thread in English criminal law as articulated in *Woolmington v. DPP*, that the State must establish guilt; regrettably, a defendant faces an uphill battle to cast doubt on machine-produced evidence. The rule's origin as a time-saving device (e.g., to avoid calling an expert for every timestamp or speedometer reading) is understandable, but extending it to today's complex, opaque software systems create an imbalance exacerbating the already assymetric structure of the criminal process. It places a technical evidential burden on, routinely, lay defendants who, understandably, lack the resources or knowledge to investigate large-scale software for hidden errors or biases. In principle the hurdle to rebut is low, but in practice it has proven unmanageably onerous – a point now acknowledged by legal experts and the courts.

Real-world experience has exposed the flaws in assuming digital evidence is infallible. The Post Office's Horizon IT system is the clearest example, as acknowledged in the Ministry of Justice's current call for evidence. Both the Post Office and the software supplier (Fujitsu) assured courts that Horizon was robust and "the computer is always right", invoking the legal presumption in their favour. Mr Justice Fraser's judgment in *Bates v Post Office Ltd (No. 6: Horizon Issues)* (2019) conclusively found that Horizon contained software defects that could cause unexplained discrepancies in branch accounts. Unable to effectively challenge the complex software, many defendants were convicted, some imprisoned, and many ruined financially. It was precisely the presumption of reliability that allowed these miscarriages to occur: earlier individual challenges by sub-postmasters failed to surmount the evidential burden to rebut Horizon's accuracy. Only a large-scale litigation with extensive expert scrutiny finally uncovered the truth, something beyond the reach of an ordinary defendant. With the increases in the uses of probabilistic systems in the criminal justice system, the injustices displayed in *Bates* can easily be repeated. In this call for evidence, there is a movement towards redressing this situation; however, this must be done with great care. Current technologies are rapidly evolving but assessing how these systems make decisions is becoming increasingly difficult.

Empirical research in computer science buttresses this point: even well-tested programs can harbour latent bugs that cause unpredictable failures. Computer software is fundamentally different from a simple mechanical clock – defects in code can remain hidden until triggered under specific conditions. As Mason observes, errors in software may produce unexpected and sometimes devastating consequences, yet may not be readily apparent to a layperson. In the Horizon case, the defendants had "no means of providing evidence to the court capable of rebutting the presumption" that Horizon was working correctly. More broadly, when a substantial institution controls the software, a defendant's challenge is frequently *insurmountable*. The barriers include lack of access to the program's source code or logs, insufficient technical expertise, and the prohibitive cost of hiring experts to interrogate complex systems. In practice, rebutting the presumption presents insuperable and costly problems for defendants. It is telling that the Law Commission's expectation of fairness in 1997 has proven wrong in hindsight.

Similar issues can occur with software systems used by governmental bodies, such as police forces. UK police forces deploy a plethora of algorithmic systems, including biometric analysis, Large Language Models (LLMs), risk prediction, and computer vision. Although there have not been known cases from the UK where the evidential reliability of these probabilistic algorithmic systems used by police forces have been questioned in court, there have been similar cases from other jurisdictions. For instance, in Ohio, a judge dismissed the use of facial recognition submitted by Cleveland Police as evidence to obtain a search warrant in a criminal case. In this case, the judge likened the facial recognition match to the statement of an anonymous informant, meaning that the software-generated output was not enough to demonstrate probable cause. Although the search of the accused's home resulted in the police finding the alleged murder weapon, this evidence was not accepted in court as the search itself was deemed to be based on inadmissible evidence. Facial recognition technologies are used frequently by UK police forces, and the human rights and equality

issues associated with disproportionate and unlawful uses of the technology have been highlighted in the *Bridges* case. With the current presumption associated with software-generated evidence, it is likely that biometric recognition technologies may be indisputably accepted as evidence for a search warrant or in a criminal case, despite the possibility that algorithmic output from such systems could harbour bias against certain races or genders.

At this point, it is important to highlight the distinction between admissibility and weight. While admissibility soft-law-criteria – like those in the Criminal Practice Directions – are essential and add flesh to the test laid down in *R v Dlugosz* ('sufficiently reliable scientific basis'), they do not opine or let alone determine whether the fact-finder will ultimately accept or reject the evidence. Indeed, juries and magistrates retain full discretion in assessing the value of admitted evidence, which makes the standardisation of reliability tests all the more vital at the point of admission. Digital evidence specific criteria could help courts assess admissibility more consistently. These include: whether the system deployed is subject to independent regulation; whether its outputs are interpretable without specialist inference; whether the inputs and processing are transparent; and whether calibration and usage are subject to routine checks.

To sum up, far from being a benign time-saver, the current presumption has fostered a systemic misunderstanding of digital evidence and lulled courts into overlooking the real possibility of software error. As such, the current presumption is not fit for purpose in modern prosecutions. It rests on outdated assumptions, unjustly reverses the burden of proof, and is difficult to challenge even when evidence is unreliable. It is therefore our position that the use of such evidence must, at a minimum, carry only a rebuttable presumption of reliability. More importantly, the burden should be on the prosecution to prove, beyond reasonable doubt, that the evidence is both reliable and appropriately generated.

2.3. Learning from other jurisdictions

As part of the scoping stage for the PROBabLE Futures project, a member of the project team conducted a research visit at Monash University in Melbourne, Australia. In this section, we discuss legal developments in Australia related to the reliability of algorithmic evidence in court, which serve as valuable examples for the UK to consider. Many regions in Australia have already assessed the need for greater legal safeguards for the use of algorithmic systems by public bodies. Notable examples include the Royal Commission inquiry into the 'Robodebt' Scheme and the guidelines on the use of Generative AI (Gen AI).

According to the Judicial Commission of New South Wales, an independent body that provides judicial education and resources, the admissibility and reliability of evidentiary outputs generated by computer algorithms require significant consideration in light of the current proliferation of artificial intelligence in the legal sphere. The Commission emphasises that it is important to acknowledge that an algorithm can only process the dataset it has been given, albeit that may be a large dataset. This is starkly different from "human intelligence" and reasoning. One type of machine learning algorithm is Gen AI. The Office of the Victorian Information Commissioner recently published guidelines on the use of Gen AI in the Victorian

Public Sector organisations, including the use of Chat GPT. The Commissioner has stressed that generative AI tools adopted by the public sector must not be used to make decisions that "may have consequences for individuals or cause them significant harm", due to the risk of inaccurate decisions and uncertain outputs. This should be an important consideration in the guidance for the admissibility of Gen AI outputs as evidence. Police forces are increasingly using Gen AI, in the forms of Natural Language Processing (NLP) and Large Language Models (LLMs). For instance, a police officer's body-worn camera could have in-built transcription technology (NLP), or a police officer could use an LLM to assist in writing a report. It is important to consider what would happen when these documents are handed over to the Crown Prosecution Service (CPS) and submitted as evidence in court. Under the current evidentiary presumption for documents produced by a "computer", it could be assumed that the Gen AI system functions correctly unless explicit evidence suggests otherwise. The Victorian Information Commissioner has emphasised that public sector organisation must establish guidelines on "how any Generative AI outputs will be assessed, valued and securely managed", including identifying any privacy impacts.

The Supreme Court of Victoria has also acknowledged the implications of using AI tools in litigation. According to the guidelines issued by the authority, if a litigant in person or a witness has used Gen AI to produce a legal document, they must provide an additional statement that clarifies the use of the algorithmic system. The Court has added that this will allow for "a more accurate assessment about the level of legal knowledge or experience possessed by a self-represented party". This is particularly relevant given the several cases, including *DPP v Khan* in Australia concerning parties submitting documentation to courts that are likely to have been prepared using Gen AI. This is evident from the misleading information or 'AI hallucinations' present in these submissions. In *DPP v Khan*, the character reference submitted was clearly a statement of irrelevant information about the defendant, indicating that it was most likely created using an LLM, which is incapable of producing human judgement about an individual in the context of the facts of the case. Judge Mossop asserted:

> *"It is clearly inappropriate that personal references used in sentencing proceedings are generated by, or with the assistance of, large language models as, if they are not objected to on that basis, it becomes difficult for the court to work out what, if any, weight can be placed upon the facts and opinions set out in them".*

There has been a similar case from England and Wales where a litigant in person submitted fake cases to the First Tier Tribunal. The Supreme Court of Victoria's guidance on Gen AI has highlighted that the outputs generated by these systems are:

> *"not the product of reasoning. Nor are they a legal research tool. They use probability to predict a given sequence of words. Output is determined by the information provided to it and is not presumed to be correct. The use of the commercial or freely available public programs such as ChatGPT and Google Gemini, is more likely to produce results that are inaccurate for the purpose of current litigation. Generative AI does not relieve the responsible legal practitioner of the need to exercise judgment and professional skill in reviewing the final product to be provided to the Court".*

This statement acknowledges that software-generated output cannot replace human judgement, such as that of a legal professional. It raises questions about how software-generated output can be taken as evidence without extensive scrutiny. This also shifts the onus of the decision back to a human expert. The guidelines from the Supreme Court assert that the use of Gen AI systems in the creation of expert reports or opinions should adhere to the Expert Witness Code of Conduct (Victoria). Building on this point, we propose that the UK guidance on digital evidence should include the requirement for an expert to demonstrate the accuracies or inaccuracies associated with algorithmic systems, including the inner workings of the technology. This is important, as algorithms have been shown to have transparency issues ('black box' problems) surrounding how they make decisions.

The Chief Justice of New South Wales has published even stricter safeguards for the use of Gen AI in affidavits, witness statements, or other evidentiary material. While the Supreme Court of Victoria has required additional documentation specifying the use of machine-learning tools for evidentiary material, the New South Wales guidelines emphasise that Gen AI must not be used in the creation or refinement of affidavits, witness statements, character references, used as evidence or during cross-examinations. Instead of requiring a statement accompanying its use, the Chief Justice of New South Wales mandates that evidentiary material be accompanied by a disclosure stating that Gen AI has not been used. The guidelines state that there may be exceptional cases where Gen AI can assist in the creation of any annexure to an affidavit, but such requests must include a statement detailing its proposed use, the specific Gen AI system and version employed, whether it is a closed-source or open-source program, privacy and confidentiality settings, and the benefits of its use for the annexure. The Chief Justice's guidelines also emphasise that Gen AI cannot be used by judges in New South Wales when creating their reasons for judgment or when assessing evidence in preparation for delivering judgments. It is evident that New South Wales has exhibited great caution regarding the outputs of similar technologies for many years, as demonstrated by the *Lie Detectors Act 1983.* In this legislation, it is emphasised that any output (or the analysis of such output) from an instrument or apparatus used to "measure or monitor physiological reactions of the body…or elements of stress, tonal variation or vibration in the voice…" is not admissible in evidence. This would cover tools like polygraphs and in recent times, controversial use cases of AI to analyse facial expressions and emotions such as the EU-funded iBorderCtrl project.

3. Ensuring Reliability: Recommendations for Reform (Question 3)

If the law is to be amended, the optimal solution is to restore robust safeguards around digital evidence so that accuracy is properly established in court, while maintaining practicality. The following proposals address the sub-questions: procedural safeguards, futureproofing, operational practicality, and the role of experts.

3.1. Placing the burden on the proponent of digital evidence

A fundamental reform would be to reverse the currently valid presumed fact and require the party relying on digital evidence (usually the prosecution in criminal cases) to demonstrate its

reliability. This suggestion aligns with what had previously been recommended by Bohm et al. In other words, the law should no longer assume that "the computer is always right" but should treat accuracy as something that must be *proven* or presumed on the basis of clearly defined presumption-raising facts adduced by the party intending to rely on digital evidence. This would effectively restructure and reinstate the principle behind former PACE s.69 in a modernised form: before digital evidence is admitted or conferred weight, there should be sufficient evidence that the system was functioning correctly and was properly used at the relevant time. Such proof might include evidence of the software's testing, error rates, security measures, and whether it was operating normally when producing the record in question. Crucially, this shifts the evidential burden back to where it belongs – on the prosecution – rather than foisting an investigative burden onto defendants who are ill-equipped to carry it. Had this been the requirement during the Post Office trials, it is likely that most prosecutions would not have succeeded, as the Post Office would have been unable to prove Horizon's reliability to the necessary standard. Indeed, experts from the IT profession have publicly called for the law to be changed in the way outlined above, to prevent future injustices. Adopting this reform would realign the law with the core tenets of criminal justice (the presumption of innocence and burden of proof on the accuser) and provide a strong incentive for institutions to vet and maintain their software systems rigorously before using their output as evidence.

3.2. Procedural safeguards: transparency, disclosure and fair testing

In addition to reviewing the presumed fact and the resulting burden of proof, specific procedures should be implemented to ensure transparency and fairness when digital evidence is used. First, there must be robust disclosure obligations regarding such evidence. The prosecution (or any party tendering digital evidence) should be required to disclose any information that might bear on the system's reliability. This includes known bugs or error reports, maintenance and update history, logs of relevant transactions, and the existence of any remote access or manual interventions in the system. In the Horizon saga, a critical failing was the Post Office's lack of transparency: for years, it withheld information about bugs and remote access that undermined Horizon's integrity. A new framework should impose a positive duty on prosecutors to investigate the reliability of their digital evidence and disclose problems, as was eventually acknowledged in the Horizon appeals: the Post Office accepted that when a case hinges on computer data, the prosecutor has a duty to assess and ensure the accuracy of that data and pursue lines of inquiry about potential flaws (see *Hamilton & Ors v Post Office Ltd* in paragraph 70, *"In relation to its duties as a private prosecutor, POL accepted that in cases where the reliability of the ARQ data was essential to the prosecution case, it had a duty to assess that data; and that in view of the limitations on the extent to which SPMs could investigate discrepancies in Horizon, POL had a duty to investigate to ensure that the evidence was accurate and to pursue reasonable lines of enquiry raised by the SPM"*). Formalising this duty would prevent "trial by ambush" where the defence is kept in the dark about software issues. It would also align with the existing prosecution duty to disclose exculpatory material (*Criminal Procedure and Investigations Act 1996*, s.3) by explicitly recognising that malfunctions in software constitute important *exculpatory evidence*.

Secondly, courts should facilitate early scrutiny of digital evidence through pre-trial case management. For instance, when a prosecution intends to rely on an automated system's output, the defence could be entitled to request a *reliability hearing* (a *voir dire*) to examine the evidence's foundation. At that stage, the court could require the prosecution to call an expert or present documentation to establish the system's soundness. The defence would have the opportunity to raise specific concerns (e.g. irregularities in the data or inconsistencies suggesting a glitch). This procedural safeguard ensures that disputes over reliability are dealt with *before* trial (or at least before the evidence goes to a jury), much as courts handle disputes over scientific evidence or confessions in a *voir dire*. By resolving reliability issues early, it streamlines the trial and protects the fairness of proceedings.

Thirdly, access to independent expertise must be enabled. In view of technological complexity, defendants should be allowed and aided to instruct independent forensic experts to examine the software or its output where feasible. In many cases, the defence will need expert assistance to identify potential malfunctions – something recognised by the consultation. If legal aid funding is a barrier, the system should accommodate requests for public funding for a defence expert when digital evidence is pivotal. A legal system which does not provide defendants with adequate funding to challenge digital evidence, should not rely on such evidence in the first place. Additionally, in cases of particular complexity, courts might consider jointly appointing an independent expert (agreed by both parties) to neutrally assess the system's reliability and report to the court. While expert testimony can be costly, it is a proportional safeguard when someone's liberty may hinge on the workings of a computer program. Even experienced IT professionals may find it challenging to detect subtle errors in large systems like Horizon; a lay defendant cannot reasonably be expected to do so without expert help. Therefore, facilitating expert involvement is essential to secure equality of arms.

Finally, auditability and record-keeping should be improved as a safeguard. Systems that generate evidence (especially those regularly used in law enforcement or commercial contexts) should maintain detailed logs and audit trails that can be produced in court. If a device or software is intended for evidential use (for example, an automated breathalyser, speed camera, or forensic software), it should be subject to rigorous quality assurance and perhaps a certification regime. A useful model is the Home Office type-approval system for speed enforcement devices which ensure these technologies are tested for accuracy and such records are available to the defence. A similar approach could be extended to critical software – e.g. requiring that any software whose output will be used in court come with documentation of testing and an attestation of its accuracy under defined conditions. These measures bolster confidence that when digital evidence is presented, it has undergone some independent verification rather than being a black box.

3.3. Futureproofing for technological advances (AI and beyond)

Any legal reform must be future proof, given the rapid evolution of technology and the increasing role of artificial intelligence in generating evidence. The next generation of cases may involve evidence produced by machine-learning algorithms – for example, facial recognition matches, predictive policing tools, or AI-driven forensics – which pose new

challenges for reliability. Unlike traditional algorithmic tools, machine-learning systems are probabilistic and adaptive, often lacking a transparent rationale for their outputs. One cannot assume they will "do as instructed" in the straightforward way a simple software might; their operation may only be understood in terms of statistical accuracy. Statistical evidence alone should not be enough to convict someone. What is more, liberal legal systems introduce the so-called specific-evidence rule preventing a legal decision relying exclusively on probabilistic or statistical evidence.

To future-proof the law, the definition of "software-generated digital evidence" should explicitly include outputs from algorithmic processes (this sub-categorisation is further discussed under *section 4.2* of this submission). The reliability inquiry should then be adapted to these contexts: for instance, requiring evidence of an AI model's validation, its error rate, and the absence of bias. If an AI system is used to produce evidence (say, an automated image comparison or voice recognition result), the proponent should demonstrate the model's performance metrics (false positives/negatives, confidence intervals) and how it was trained or calibrated. Courts may need to hear from data scientists to understand whether an AI's output in a given case can be trusted. In short, the same principle of caution must apply – the more complex and novel the technology, the greater the need to *prove* its reliability rather than assume it.

To keep pace with innovation, the framework should be technology-neutral to the extent possible. Rather than enumerating specific technologies, it could impose general criteria (e.g. that any algorithmic evidence must be shown to be operating as intended, and any limitations or error margins disclosed). This can be supported by guidelines or practice directions that can be updated more easily than primary legislation. Additionally, continuous monitoring is wise: a standing advisory group of judges, scientists, and technologists could periodically review how the rules are functioning with new tech (such as AI, blockchain evidence, Internet-of-Things sensors, etc.) and recommend adjustments. The goal is to avoid a scenario where the law reforms fix yesterday's problem (like Horizon) but become obsolete with tomorrow's technology. By embedding flexibility and a focus on core principles of reliability and transparency, the regime will be resilient against future changes. As Marion Oswald noted in the context of AI/ML: *"Forecasts, classifications or predictions produced by many existing algorithmic tools are probabilities…"* The law must therefore be prepared to interrogate and doubt AI outputs just as (if not more than) other digital evidence.

3.4. Ensuring operational practicality

The proposed solutions, while stronger, must remain practicable in the criminal justice system. A concern might be that requiring proof of reliability for every piece of digital evidence could burden courts and prosecutors. This can be mitigated by scoping and case management. Firstly, as hinted under Question 4 of the consultation document, the reform should target a narrow category of evidence – *evidence generated by software or algorithms* – not all digital evidence at large (further discussed under *section 4.2* of this submission). This means straightforward digital records (emails, documents, CCTV videos, etc.) would not trigger a reliability proof requirement, only those where the fact at issue is *produced* by

software processing. This also extends to widely accepted digital forensics technologies, such as a breathalyser. Williams has argued that in drink-drive cases, the defence can request "copious" amounts of documents from the prosecution on how an alcohol reading was calibrated by the breathalyser device, "in the underlying hope that these are not produced". Here, the burden of proof lies on the prosecution. However, in the case of a breathalyser, the digital evidence is produced through a much simpler calibration than an algorithm. This device analyses the Blood Alcohol Concentration (BAC) in the expired air of an individual, using electrochemical or infrared technology. Thus, the process behind generating the output can be made clear through documentation provided by the manufacturer or certification body, without 'black-box' issues. On the other hand, with probabilistic technologies (such as tools used for identifying individuals as high risk or measuring emotional states), the number of inputs and outputs may vary; as such, there may be no objective way to verify such outputs – the absence of ground truth. In addition, the type of output might also vary; for instance, the output could be a binary number or a likelihood scale. By focusing on cases where the computer output is pivotal and essentially substituting for a human witness (for example, an algorithm identifying a suspect), we limit the number of instances where detailed proof is needed. Many prosecutions will continue unaffected, as they do not rely solely on machine-generated facts. It is also evident from the Manchester drink driving cases (see *R. v Bolton Magistrates' Court* [1991] and the Randox Testing Services scandal) that even accepted scientific and objective evidence submitted by prosecution requires further proof.

Secondly, even within this category, the burden of proving reliability need not be onerous if the system is genuinely reliable. Often a short witness statement or certificate from a system maintainer may suffice, unless the defence raises specific challenges. Only when a real question of malfunction arise would a deeper inquiry be needed. A similar approach can be adopted: routine evidence from computers can be accompanied by a basic assurance of proper operation (subject to penalty if false), and only contested or high-stakes instances would escalate to requiring expert evidence. In essence, the law can expect parties to *be prepared* to establish reliability, but not every case will require a full-blown "trial within a trial" on the software. Judges should have discretion to tailor the extent of scrutiny to the needs of the case – for instance, dispensing with formal proof if the defence does not dispute the evidence, or conversely, ordering a thorough examination if the evidence is both crucial and credibly questioned.

It is also important to leverage existing technical standards to aid practicality. Where independent certifications or audits of a system exist, these can be used in court to streamline proof. For example, if a forensic software tool adheres to an industry standard (perhaps ISO/IEC standards for digital evidence tools) and has documented test results, that can be presented as part of the reliability foundation. Similarly, if a device has been approved by a regulatory body (e.g. the Home Office Science unit approving an evidential breathalyser model), a certificate or expert report to that effect could be taken as prima facie evidence of reliability – unless the defence shows reason to doubt it. This approach balances efficiency with the defendant's right to challenge: most of the time a certified system will not be challenged, but the door remains open for challenge if something specific seems awry.

Finally, concerning expert witnesses (Question 3(d)): the legal system will increasingly need access to *specialist expertise* in software engineering, digital forensics, and related fields. At present, the question arises whether sufficient experts exist and how to ensure quality. There are current shortcomings in regulation and accreditation, particularly for emerging technologies. Noteworthy is the limited scope and patchy compliance with existing regimes like those under the Forensic Science Regulator (See, Peter Sommer, Angus Marshall, and Emmanuel Amoako where they individually provided important insights into the difficulties of regulating digital forensic practices). Even where standards like ISO exist, the rapid development of digital technologies and their diversity make it difficult for both prosecution and defence experts to maintain compliance. This raises concerns about creating a system that inadvertently penalises the defence, especially where expert witnesses cannot meet accreditation standards due to the lack of available insurance or formal recognition of new methods. Therefore, a fair balance must be struck: while evidence used by the prosecution should meet high regulatory and accreditation standards, the defence must retain the ability to challenge such evidence without being unfairly barred on technical grounds.

In the short term, courts can utilise established expert registers (for example, the National Crime Agency's Expert Advisers Database (EAD), or independent expert witness directories such as the UK Register of Expert Witnesses) to identify qualified individuals. Over time, investment in training and accreditation in the field of digital evidence should be encouraged. If demand increases due to these legal changes, the market and professional bodies (like BCS, The Chartered Institute for IT, or the Academy of Experts) are likely to respond by training more specialists in this niche. It may also be worth considering joint instruction of a single expert in cases where the basic integrity of a system is in issue, to avoid "battle of the experts" and reduce costs. In summary, while resource challenges exist, they are manageable with forward planning. The integrity of verdicts must take priority over convenience; as Dr. Sam De Silva (Chair of BCS's Law Specialist Group) aptly noted, expecting non-IT-specialist defendants to prove a complex system wrong is untenable – even seasoned IT professionals would struggle. It is more practical and just for the system to supply the necessary expertise to examine the evidence, rather than to assume infallibility and risk injustice.

3.5. Ensuring reliability through a robust (probabilistic) framework for evaluative reporting

What is more, the topic of digital evidence needs to be nested within and examined in the context of forensic science and the deeply problematic, arguably dysfunctional UK forensic ecosystem especially following the axing of the Forensic Science Service in 2012. For the probative value of digital evidence depends on, and is heavily shaped by, the underlying methodology followed by digital evidence examiners and the testimonial claims made by expert witnesses. More specifically, the significance of digital evidence can be reduced to the evaluation of such evidence. Consumers of digital evidence, particularly the judiciary, would therefore be well advised to keep a critical eye on how the discipline deals with these challenges. In this context one would expect the field of digital evidence to demonstrate a high level of understanding of, and adherence to, sound evaluation principles. Regrettably, both the digital evidence literature and the field are not safe havens in this regard. They are rather a fertile ground for misconceptions, ad-hoc theories and ideas that pose impediments

to the implementation of coherent evaluative reporting schemes. A major concern is that evaluative reports in the digital evidence field do not deploy any probabilistic framework for measuring uncertainty. As Biedermann and Kotsoglou explain *"[w]hat this means for digital forensic science is potentially far-reaching: as a currently developing new branch of forensic science, it has a unique opportunity not to commit the failures and shortcomings in evidence interpretation that (continue to) affect traditional forensic disciplines. However, this would require the community of researchers and practitioners to draw suitable conclusions from the principles of forensic interpretation that have been developed since the middle of the last century".* Ignoring or refusing to apply a probabilistic (scientific) framework for evaluative reporting is not conducive of such a perspective.

The legal system needs to introduce specific requirements for the generation and admissibility of digital evidence, negating thus what Biedermann and Kotsoglou critically describe as 'digital evidence exceptionalism', i.e. the rather mainstream view that digital evidence is exempt from principles and methodological rules salient in every scientific discipline. The claim that digital forensic science is inherently different from other forensic branches which abide by scientific principles does not meet the normative requirements, both methodological and procedural in nature, which forensic fields ought to fulfil. A set of descriptors or unjustified conclusions that lack an underpinning conceptual framework is sidestepping the requirement for justification. A lack of accountability represents a regrettable and deleterious loss, as it removes justifications and reasoning processes entirely from the public arena especially in the critical context of criminal law. If this is what (digital) forensic science is or aims at, then it is difficult to see how it can meaningfully serve the needs of fact-finders in the pursuit of justice.

For it is vital to ensure that the probative value of digital evidence is assessed using logically sound principles, notably a Bayesian probabilistic framework (likelihood ratio) as strongly suggested by the ENFSI Guideline on evaluative reporting. The persistence of digital evidence 'experts' in not using such framework can only exacerbate already existing problems.

4. Defining "Computer Evidence" and Scope of Reform (Question 4)

4.1. Scope: Evidence generated by software vs merely recorded data

It is critical to delineate what types of digital evidence should fall within the ambit of these reforms. The consultation rightly draws a line between evidence that is "generated by software" and digital evidence that is merely captured or stored by a device. We agree that the focus of any new rule should be on the former category – i.e. information that owes its existence to a computer program's operations or analysis – and that purely recorded data (where the device is acting as a passive receptacle for human input or environmental information) should remain out of scope. This distinction ensures we target the problem (software reliability) without overburdening the handling of ordinary digital evidence. It is also important to consider that evidence that is captured by a device, such as photographs or video footage, can also be inputted into software to generate further evidence. For instance, many police forces are deploying computer vision technologies to filter through large

collections of video footage in investigations. For this, algorithmic Video Content Analysis (VCA) technologies are used to identify objects of interest, such as firearms or vehicles, in video footage from CCTV cameras or body worn cameras. In this case, the original video footage can be considered ordinary digital evidence, and the probabilistic output of the computer vision technology would be evidence generated by software.

As rightly pointed out in the consultation document, an automated accounting report like the Horizon branch accounts is clearly within scope: the figures and alleged shortfalls were the result of computations by the system's software. We also support the inclusion of other examples envisaged in the consultation document such as fraud detection software outputs (flags or scores generated by algorithms) and plagiarism detection results – essentially, any scenario where a computer program processes data and produces an assertion that is used as evidence. In addition, we propose the inclusion of machine-generated audit trails or logs (for instance, logs from a secure door entry system attributing entries to a keycard, or metadata generated by a system). Likewise, an AI-driven facial recognition match, VCA or an automated number-plate recognition hit would be in scope, as the evidence presented (a match confidence, etc.) is generated by an algorithm and is probabilistic. In all these cases, the trustworthiness of the evidence is directly tied to the correct functioning of the software.

On the other hand, one can think of digital photographs, videos, or audio recordings captured on devices – these would be out of scope of this specific reliability presumption reform (though of course still subject to authenticity rules). A CCTV camera recording an incident, a smartphone recording a conversation, or body-camera footage are examples of digital evidence that is *captured* rather than generated. The reliability question relates to the question of whether the file has been altered, or the device tampered with, not whether the device's internal software correctly performed some complex analysis. Similarly, text messages, emails, and social media posts are digital evidence but not "software-generated" in the sense at issue – the content is authored by humans; the device merely transmits or stores it. These forms of evidence raise other issues (authenticity, hearsay, privacy) but not the algorithmic-reliability-concern salient in Horizon-type cases. Including them in a new reliability rule would unnecessarily broaden the scope and could impede prosecutions with burdensome checks on mundane evidence. Therefore, the reform should be *carefully limited*: it might apply, for instance, to evidence of facts or conclusions produced by the operation of software or an automated system and explicitly exclude evidence that is only recorded or stored digitally without material transformation by the device. This approach was hinted at in the consultation document and is eminently sensible. It will ensure that the courts do not have to, say, certify that a mobile phone was working properly every time a text message is introduced, while still covering the problematic scenarios where a computer-generated result (not directly verifiable by human senses) is taken as proof of guilt.

4.2. Adopting the term "Digital Evidence" over "Computer Evidence."

The question also asks how "computer evidence" should be defined for these purposes. In answering that, it is worth addressing terminology. The law has historically referred to "computer evidence," as in PACE s.69 and related case law, but this term is increasingly seen

as outdated. Modern jurisprudence and academic discourse prefer the broader term "digital evidence" (NB: this is the term adopted in this submission). Digital evidence is commonly defined as any "*information of probative value that is stored or transmitted in binary form*". This encompasses not only traditional computers, but also smartphones, tablets, servers, cloud databases, IoT devices, and all manner of electronic systems. Using the term "digital evidence" acknowledges that we live in a ubiquitously digital world: a car's sat-nav, a smart doorbell's video, a smart-fridge or washing machine, or an online transaction record are all digital evidence, even if no one would colloquially call some of those things "computers." In contrast, the phrase "computer evidence" might misleadingly imply a narrower scope (perhaps one might think of a desktop PC). As the consultation document itself notes, what counts as digital material has expanded enormously since the 1980s. It would be prudent for any new legal provisions to use up-to-date terminology that covers the full range of relevant technology.

In fact, the shift in terminology is reflected in expert literature and practice. For example, there are dedicated law journals titled *FSI: Digital Investigation or Digital Evidence and Electronic Signature Law Review*, indicating how the field conceptualises this area of evidence. Courts and practitioners internationally speak of electronic or digital evidence as a category. The term "digital" also aligns with other initiatives, such as the Council of Europe's Electronic Evidence Guide, and standards by groups like the Scientific Working Group on Digital Evidence (SWGDE) – whose definition we quoted above. Adopting "digital evidence" would harmonise our language with these global standards and avoid confusion. It is also more future-proof: fifty years from now, the devices producing evidence may not resemble "computers" as we traditionally have known them, but they will in all likelihood still produce digital data of some form. The law should capture the essence (the data is digitally produced) rather than the form (the device is a 'computer').

In view of that, within the broad universe of digital evidence, we still need a term for the subset this reform concerns. We might refer to it as "software-generated digital evidence" for precision – meaning digital evidence that is the direct output of a computer process. The statute or guidance could define this term. For instance: *"software-generated digital evidence" means information intended to be relied upon as evidence of a fact, which information was produced by the application of computer software or an automated algorithm, rather than by direct human perception or input."* This would firmly capture things like automated analysis, system-generated logs, and algorithmic results. Meanwhile, it would exclude things like a written Word document (the computer helped type it, but the content is human), or a scanned copy of a paper.

In more concrete terms, we propose a three-tier hierarchical framework (see *figure 1, page 1*) to ensure that the reformed rules on reliability and rebuttable presumptions are confined to the appropriate context. Guiding this hierarchy are three core criteria. First is complexity, ranging from probabilistic AI outputs (multiple inputs, dynamic decision-making) to deterministic devices (single-input, rule-based results). Second is human interaction, i.e., the degree interpretation required, whether AI merely flags data for human review or operates

autonomously. Finally, is the issue of regulatory backing, from independent certifications to manufacturer self-assessment, shaping the burden of proof.

At the apex of this hierarchy would be highly complex, probabilistic systems – those incorporating AI and machine learning – where outcomes are not easily understood or verifiable, and which often rely on opaque or proprietary methodologies. These systems, (including biometric recognition algorithms, generative AI outputs, or risk-assessment models) often with multiple inputs, probabilistic processing, and inferential outputs, should be subject to rigorous reliability testing. due to opaque decision-making, multiple inputs/outputs, and human interpretation requirements. For admissibility, prosecutors should be required to furnish exhaustive documentation: training data sources, error rates across demographic groups, and peer-reviewed validation studies. Independent accreditation (e.g., ISO standards specific to algorithmic fairness) and expert testimony from data scientists would further ensure accountability.

The middle tier would encompass semi-automated technologies, such as AI-assisted video review (e.g., systems flagging drivers for phone use), GPS data integrated with speed cameras, and DNA mixture analysis. While these tools may appear objective, their outputs often mask interpretive complexities. For instance, a system might flag a driver's hand movement as phone use, oblivious to contextual nuances like Bluetooth connectivity. To mitigate risks, this tier warrants a rebuttable presumption of reliability. Prosecutors must first demonstrate adherence to manufacturer protocols (e.g., calibration records for traffic cameras), while defendants retain the right to challenge evidence through alternative data or expert critiques of systemic limitations. For example, in 2020, Peter Marrable was summoned to court for driving at 72mph in a temporary 50mph zone. Using GPS data, Mr. Marrable was able to successfully rebut the speed camera's accuracy.

At its base, the hierarchy would include regulated, rule-based devices like breathalysers or body cameras, which operate on fixed parameters with minimal human interpretation. While these systems are simpler (and often with binary or easily understood outputs), they still warrant caution and should also not be subject to absolute presumptions of reliability. Although devices like breathalysers benefit from routine inspections (e.g., UKAS audits), it is important that the defence have access to calibration records. In other words, while prosecutors need only prove operation per manufacturer guidelines (e.g., calibration certificates), defendants must retain the right to demand maintenance records or challenge evidence through independent testing.

To conclude on this point, embracing the term "digital evidence" in our law would be a positive modernisation supported by both technical understanding and comparative jurisprudence. It underscores that evidence from computers are part of a larger continuum of electronic evidence, and it prepares our legal language for whatever new devices and systems the future may hold. Within that, precisely defining the subset of digital evidence that triggers heightened reliability scrutiny will guard against overreach. Yes, evidence generated by software should squarely be in scope (with examples as discussed), and evidence merely recorded by devices (with no complex transformation) should be out of

scope – this delineation is logical, fair, and supported by the rationale of the reform. More importantly, the hierarchy model proposed would need flexibility and continued refining, especially in the vast and shifting "middle tier" between simple and complex systems. Implementation should therefore include a case-by-case assessment mechanism informed by the following clear principles: regulatory compliance, interpretability, calibration, human interaction, and independent scrutiny.

5. Additional Considerations (Question 5)

Beyond the core issues above, there are other important factors to consider in reforming the use of digital evidence:

5.1. Judicial training and understanding

The courts must be equipped to engage with technical evidence effectively. The Horizon saga revealed that misunderstandings about technology can be pervasive. In Christie's words, early on, some judges *"…almost seemed to take pride in the ignorant waffle about computers that they airily spouted from the bench"*. It is crucial that judges (and lawyers) receive training on digital evidence principles, including the inherent fallibility of software and how to handle expert testimony about it. Improved digital literacy in the judiciary will help ensure that any new rules are applied rigorously and intelligently, rather than treated as a mere formality.

5.2. Guidance on evaluating digital evidence

Hand in hand with training, it would be useful to develop judicial guidance or a Practice Direction on factors to weigh when assessing software reliability. This could include a non-exhaustive list of considerations (e.g. whether the software was developed to industry standards, whether it has been independently audited, whether the particular output falls within its normal operating parameters, known error rates, etc.). An example of an evaluation tool for probabilistic AI systems is a checklist we are creating as part of the PROBabLE Futures project. This checklist is a living document that is being drafted in conjunction with the National Police Chiefs' Council (NPCC) for the assessment of AI (and other emerging data analytics tools) used in law enforcement.[2] Such guidance would aid consistency across cases and give parties a clearer idea of what is expected if they seek to rely on, or to challenge, digital evidence.

5.3. Disclosure and inter-partes cooperation

As already noted, much of the difficulty for defendants arises from information asymmetry – the prosecution or a third-party institution holds all the cards about the system. To mitigate this, the disclosure regime must be enforced strictly. Courts should be ready to sanction non-disclosure of relevant digital evidence (for instance, by staying proceedings or excluding evidence if the prosecution fails to disclose a known issue or blocks reasonable defence

[2] For more information on this checklist, please contact the project team.

inquiries). In complex cases, courts could also encourage cooperation between opposing experts to narrow down points of disagreement about the technology, which can save time and avoid technical obfuscation.

5.4. Protection of sensitive software (trade secrets/national security)

One oft-cited challenge is that software might be proprietary or sensitive – dubbed "source code secrecy" (for example, source code owned by a company or government). While intellectual property or security concerns are valid, they cannot be allowed to trump a defendant's right to a fair trial. Mechanisms exist to handle this tension – for instance, courts can order disclosure of code under strict confidentiality undertakings, or review by a trusted independent expert who can report on relevant issues without revealing trade secrets. In camera proceedings or confidentiality rings can protect genuinely sensitive information. The key point is that such concerns must not serve as a blanket excuse to withhold evidence of potential faults. If a system's inner workings are so secret that even a controlled examination is impossible, one must question whether its outputs are suitable for use in criminal court at all. In short, fair trial rights should take priority, with creative solutions employed to address secrecy where necessary.

5.5. Resource implications

Implementing these changes will require resources – notably, the time of experts and judges. It is important that the Ministry of Justice and other stakeholders plan for this. Legal aid budgets may need adjustment to fund expert challenges in appropriate cases; prosecutors may need training and perhaps additional expert support to certify software evidence upfront. For instance, in response to the concerns raised by the Law Commission in 2011 on the need for reliable expert evidence and additional pre-trial hearings, the Government at the time stated that this would result in "additional costs, but without sufficient reliably predictive savings to compensate for those costs". However, these costs are justified when weighed against the costs of wrongful convictions (human, financial, and reputational costs). The Post Office scandal itself has led to substantial compensation bills and a public inquiry – a stark reminder that unsafe convictions are far more costly in the long run than doing justice right the first time.

5.6. Monitoring and continuous improvement

Finally, it would be wise to monitor the impact of any reform. For example, after a few years of the new regime, data could be gathered on how often digital evidence is being challenged, how often issues are found, and whether certain agencies have recurring problems. This can inform whether further measures are needed (or conversely, if the system is working well, it can reassure that the balance is right). The law in this area should remain responsive to experience, much as it is doing now by reviewing the presumption after the Horizon experience.

6. Conclusion

To summarise our submission, the current presumption that digital evidence is reliable unless disproven is ill-suited to today's digital justice system and has been shown to undermine fairness. Whilst reverse burden of proof clauses are not per se unlawful, this particular one is not only unmanageable for the defendant but deleterious for the legal system. The recent miscarriages of justice in the context of, arguably, the biggest legal scandal in the history of the UK, are merely a manifestation of that structural problem. The technological advances today, facilitated by machine-learning algorithms, go far beyond the "computer records" and documents originally conceived in the PACE. These advancements were also not readily available in 2000, when the Law Commission's presumption was introduced. Adopting the term "digital evidence" in place of "computer evidence" will reflect the modern reality and ensure the law uses accurate, forward-looking language. We have suggested a three-tier hierarchal framework (Figure 1) to demonstrate the different types of digital evidence: complex probabilistic software systems; semi-automated technologies; and simpler rule-based devices. We have also suggested that "software-generated digital evidence" could be defined as *"information intended to be relied upon as evidence of a fact, which information was produced by the application of computer software or an automated algorithm, rather than by direct human perception or input."*

Reforms are necessary to restore balance: the prosecution should bear the burden of establishing software reliability, and robust procedural safeguards (early disclosure, expert involvement, and judicial oversight) should be in place to allow defendants a fair opportunity to challenge algorithmic evidence. Any new framework must be carefully scoped to target digital evidence of high-risk nature and to remain adaptable as technology evolves, especially with the advent of AI-driven tools. In the creation of these guidelines, lessons must be learned from caselaw and empirical studies, including from jurisdictions that have clear guidance on the admissibility of software-generated evidence, including Australia.

These changes, taken together, will promote transparency, accuracy, and justice in criminal proceedings. The legal system cannot afford to place blind faith in computers – it must insist on demonstrable trustworthiness before depriving someone of their liberty based on digital evidence. The experiences and sources cited in this submission strongly support this approach, which will ultimately strengthen confidence in our courts' ability to handle technology while safeguarding the rights of all parties.

Bibliography

Bates v The Post Office Ltd (No 6: Horizon Issues) Rev 1 [2019] EWHC 3408 (QB)

Biedermann, A. and Kotsoglou, K. (2020) 'Digital Evidence Exceptionalism? A Review and Discussion of Conceptual Hurdles in Digital Evidence Transformation' *Forensic Science International: Synergy* 2: 262-274, https://doi.org/10.1016/j.fsisyn.2020.08.004

Bittle J. (2020) 'Lie Detectors have Always Been Suspect. AI has Made the Problem Worse', *MIT Technology Review*, 13 March. Available at: https://www.technologyreview.com/2020/03/13/905323/ai-lie-detectors-polygraph-silent-talker-iborderctrl-converus-neuroid/ (accessed 9 March 2025)

Bohm *et al.* (2022) 'Briefing Note: The Legal Rule That Computers are Presumed to be Operating Correctly – Unforeseen and Unjust Consequences' *Digital Evidence and Electronic Signature Law Review* 19: 123-127, https://doi.org/10.14296/deeslr.v19i0.5476

Carter, Z., Radcliff, E. and Dowel, G. (2024) 'When to Put a Ring on it? Confidentiality in Litigation Proceedings' *Macfarlanes*, 27 August. Available at: https://www.macfarlanes.com/what-we-think/102eli5/when-to-put-a-ring-on-it-confidentiality-in-litigation-proceedings-102jhht/ (accessed 6 March 2025)

Christie, J. (2023) 'The Law Commission and Section 69 of the Police and Criminal Evidence Act 1984', *Digital Evidence and Electronic Signature Law Review* 20: 62-95, https://doi.org/10.14296/deeslr.v20i.5642

Council of Europe. (2022) 'iPROCEEDS-2: Launching of the Electronic Evidence Guide v.3.0', 23 June, Available at: https://www.coe.int/en/web/cybercrime/-/iproceeds-2-launching-of-the-electronic-evidence-guide-v-3-0 (accessed 25 March 2025)

Criminal Practice Directions 2023

Criminal Procedure and Investigations Act 1996

Cross, M. (2024) 'IT Experts Call for Review of "Computer is Always Right"', *The Law Society Gazette*, 15 January. Available at: https://www.lawgazette.co.uk/law/it-experts-call-for-review-of-computer-is-always-right/5118414 (accessed 15 March 2025)

Daniel, L. (2025) 'Judge Throws Out Facial Recognition Evidence In Murder Case', *Forbes*, 29 January. Available at: https://www.forbes.com/sites/larsdaniel/2025/01/29/judge-throws-out-facial-recognition-evidence-in-murder-case/ (accessed 11 March 2025)

Daprile, L. (2025) 'Cleveland Police used AI to Justify a Search Warrant. It had Derailed a Murder Case', *Cleveland*, 25 January. Available at:

https://www.cleveland.com/news/2025/01/cleveland-police-used-ai-to-justify-a-search-warrant-it-has-derailed-a-murder-case.html (accessed 18 March 2025)

DPP v Khan [2024] ACTSC 19

Farrands, D. (2022) 'Artificial Intelligence and Litigation – Future Possibilities', *Judicial Commission of New South Wales*, September. Available at: https://www.judcom.nsw.gov.au/publications/benchbks/judicial_officers/artificial_intelligence_and_litigation.html (accessed 12 March 2025)

Hamilton & Ors v Post Office Ltd [2021] EWCA Crim 577

Hern, A. (2024) 'Update Law on Computer Evidence to Avoid Horizon Repeat, Ministers Urged', *The Guardian*, 12 January. Available at: https://www.theguardian.com/uk-news/2024/jan/12/update-law-on-computer-evidence-to-avoid-horizon-repeat-ministers-urged#:~:text=Stephen%20Mason%2C%20a%20barrister%20and,them%20who%20has%20the%20information.%E2%80%9D (accessed 10 March 2025)

Holmes, C. (2023) 'Royal Commission into the Robodebt Scheme' 7 July. Available at: https://robodebt.royalcommission.gov.au/publications/report (accessed 10 March 2025)

Home Office. (2023) 'Guidance: Home Office Type Approved Evidential Breath Alcohol Analysis Instruments' 27 January. Available at: https://www.gov.uk/government/publications/approved-breath-alcohol-analysis-testing-devices (accessed 3 March 2025)

Home Office. (2024) 'Home Office Type Approval of Road Traffic Law Enforcement Devices: Submission Process Guidance' October. Available at: https://assets.publishing.service.gov.uk/media/672e1e703b601d048796ae54/HOTA+RTLED+Submission+Guidance+v3.0__2_.pdf (accessed 1 March 2025)

International Organisation for Standardisation. (2012) 'Information Technology — Security Techniques — Guidelines for Identification, Collection, Acquisition and Preservation of Digital Evidence', Doc ID ISO/IEC 27037:2012, Available at: https://www.iso.org/standard/44381.html#:~:text=ISO%2FIEC%2027037%3A2012%20provides,can%20be%20of%20evidential%20value (accessed 4 March 2025)

Katyal, S. (2019) 'The Paradox of Source Code Secrecy' *Cornell Law Review* 104: 1183-1279

Kotsoglou, K., and McCartney, C. (2021) 'To the Exclusion of all Others? DNA Profile and Transfer Mechanics—R v Jones (William Francis) [2020] EWCA Crim 1021 (03 Aug 2020)' *The International Journal of Evidence & Proof* 25: 135-140, https://doi.org/10.1177/13657127211002288

Ladkin *et al.*, (2020) 'The Law Commission Presumption Concerning the Dependability of Computer Evidence' *Digital Evidence and Electronic Signature Law Review* 17: 1-14 https://doi.org/10.14296/deeslr.v17i0.5143

Law Commission. (2011) 'Law Com No 325: Expert Evidence in Criminal Proceedings in England and Wales', 21 March, Available at: https://assets.publishing.service.gov.uk/media/5a7cc0c0e5274a38e575689b/0829.pdf (accessed 5 March 2025)

Lindenmuth, K. (2019) 'Prevention or Self-Fulfilling Prophecy? Predictive Policing's Erosion of the Presumption of Innocence Erosion of the Presumption of Innocence' *Law School Student Scholarship* 1018: 1-28

Mason, S. and Seng, D. (2021) *Electronic Evidence and Electronic Signatures.* 5th edn. London: University of London Press

Mason, S. (2024) 'The Presumption that Computers are Reliable', Counsel, 10 July. Available at: https://www.counselmagazine.co.uk/articles/the-presumption-that-computers-are-reliable#:~:text=The%20effects%20of%20bugs%20in,and%20costly%20problems%20for%20defendants (accessed 8 March 2025)

Marshall, P. (2020) 'The Harm that Judges do – Misunderstanding Computer Evidence: Mr Castleton's Story', *Digital Evidence and Electronic Signature Law Review* 17: 25-48 https://doi.org/10.14296/deeslr.v17i0.5172

Ministry of Justice. (2013) 'The Government's response to the Law Commission report: "Expert evidence in criminal proceedings in England and Wales" (Law Com No 325)', 21 November. Available at: https://www.gov.uk/government/publications/government-response-to-law-commission-report-on-expert-evidence (accessed 5 April 2025)

Nelson, L. (2024) 'Hallucinatory Judgments and Automated Vehicles: AI and the Law' *Deka Chambers,* 7 March. Available at: https://www.dekachambers.com/2024/03/07/hallucinatory-judgments-and-automated-vehicles-ai-and-the-law/ (accessed 10 March 2025)

Oswald, M. (2020) 'Technologies in the twilight zone: early lie detectors, machine learning and reformist legal realism', *International Review of Law, Computers and Technology* 34(2): 214-231 https://doi.org/10.1080/13600869.2020.1733758

Police and Criminal Evidence Act 1984

R (Bridges) v CC South Wales Police [2020] EWCA Civ 1058

R v Dlugosz [2013] EWCA Crim 2

R v Bolton Justices ex parte Scally and others [1991] 1 QB 537

Stoykova, R (A). (2021) 'Digital Evidence: Unaddressed Threats to Fairness and the Presumption of Innocence' *Computer Law & Security Review* 42: 1-20

Tamascelli *et al*. (2024) 'Artificial Intelligence for Safety and Reliability: A Descriptive, Bibliometric and Interpretative Review on Machine Learning' *Journal of Loss Prevention in the Process Industries* 90: 1-19, https://doi.org/10.1016/j.jlp.2024.105343

Williams, P M. (2018) 'Commentary: Current Defence Strategies in Some Contested Drink-Drive Prosecutions: is it Now Time for Some Additional Statutory Assumptions? *Forensic Science International* 293: e5-e9, https://doi.org/10.1016/j.forsciint.2018.09.030

Woolmington v DPP [1935] AC 462