Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/CLSR

Computer Law & Security Review

Digital evidence: Unaddressed threats to fairness and the presumption of innocence

Radina Stoykova [a,b,*]

[a] *Department of Transboundary Legal Studies, Faculty of Law, University of Groningen, The Netherlands*
[b] *Norwegian University of Science and Technology, Gjøvik, Norway*

ARTICLE INFO

Keywords:
Presumption of innocence
Digital evidence
Reliability
Digital forensics
Fair trial

ABSTRACT

Contemporary criminal investigation assisted by computing technology imposes challenges to the right to a fair trial and the scientific validity of digital evidence. This paper identifies three categories of unaddressed threats to fairness and the presumption of innocence during investigations – (i) the inappropriate and inconsistent use of technology; (ii) old procedural guarantees, which are not adapted to contemporary digital evidence processes and services; (iii) and the lack of reliability testing in digital forensics practice. Further, the solutions that have been suggested to overcome these issues are critically reviewed to identify their shortcomings. Ultimately, the paper argues for the need of legislative intervention and enforcement of standards and validation procedures for digital evidence in order to protect innocent suspects and all parties in the criminal proceedings from the negative consequences of technology-assisted investigations.

1. Introduction

A recent report by the NCPP in the UK points out that 90% of criminal investigations nowadays have a digital element, and identifies standardization and automation as core principles for further development of digital forensics (T. F. The UK National Police Chiefs Council 2020). Digitalization changes the methods, techniques, and the scope of criminal investigations. Despite the strive for automation, digital forensics (DF) examination still struggles with limited resources, overreliance on tools and subjective opinions. Digital evidence is increasingly presented and accepted in courts without scientific validation of the digital forensic methodology or tools. While classical investigative measures are subject to strict limits and fair trial guarantees, digital investigations still lack quality assurance and accountability. There are no European minimum standards for digital evidence to establish and enforce scientific validation in digital forensics.

Inappropriate use of poorly tested technology undermines the right to a fair trial, as formulated in Art.6 ECHR,[1] and threatens the presumption of innocence at an early stage of an investigation. Moreover, ineffective pre-trial and trial guarantees for defendants are not suitable to validate complex DF methodology and tools and the suspect/defendant position to collect or challenge digital evidence in the criminal proceedings is weak. Overreliance on and inappropriate use of technology in combination with the weak position of suspects/defendants can lead to unequal treatment of sus-

* Correspondence to: Norwegian University of Science and Technology, Gjøvik, Norway.
E-mail address: radina.r.stoykova@ntnu.no

[1] Council of Europe, European Convention for the Protection of Human Rights and Fundamental Freedoms, as amended by Protocols Nos. 11 and 14, 4 November 1950, Article 6 - Right to a fair trial.

https://doi.org/10.1016/j.clsr.2021.105575

pects/defendants and lack of legal certainty in the judicial process.

Therefore, this paper is an attempt to clarify the connection between the right to a fair trial, in particular the presumption of innocence (PI), with digital evidence rules development as a theoretical framework. Examined is to what extend digital evidence practices comply with fair trial principles and how technology-assisted investigations challenge criminal procedure. Digital evidence is defined as: any information processed by electronic medium which supports or refutes a hypothesis about the state of digital artefacts or digital events, of potential relevance and probative value for a criminal investigation[2]. Digital evidence is the result of scientific methodologies and tools which ensures that "*its authenticity and integrity can be validated*" (Casey, 2007; Mocas, 2004). Just like any other forensic science, digital forensics must account for tool, methodology and human limitations. Due to the dynamics in the field and the broad scope of application at every stage of the criminal procedure, digital forensics practice must be guided by minimum quality standards and fair trial safeguards irrespective of jurisdictional differences.

This paper classifies the unaddressed threats to PI with respect to technology-assisted investigations and digital evidence in three groups: inappropriate and inconsistent use of technology (*Section B*); old procedural guarantees, which are not adapted to contemporary digital evidence processes and services (*Section C*); and the lack of reliability testing in digital forensics practices (*Section D*). It is argued that the use of technology for investigation purposes must be evaluated against fair trial and reliability standards. Outlined are issues with *de facto* reverse burden of proof, low quality data processing, reliance on untested digital expert evidence (opinion), and lack of criminal procedure guarantees in data retention, crime prevention and suspicion-based procedures. *Section E* examines, critically, proposed approaches to address procedural fairness threats in investigations and their applicability in digital evidence context.

It is important to discuss which stage of the investigation carries high risks to the burden and standard of proof and requires PI safeguards reinforcement. Two of the four threats to the presumption of innocence as formulated by *Ashworth* (Ashworth, 2006) – *confinement* (defining offences as so to reduce the impact of PI) and *erosion* (recognizing more exceptions to PI) – are broadly discussed in the literature and further summarized with regard to surveillance and anti-terrorism legislation. However, the other two dangerous practices – *sidestepping* (imposing restrictions on the rights of non-convicted persons) and *evasion* (introducing civil law procedures in order to circumvent the rights conferred on an accused person) – have more subtle impacts on PI, because they could be justified by the need to collect evidence for the investigation. It is further argued that side-stepping and evasion practices are facilitated by the extended use of technology and the increased reliance on digital forensics in criminal investigations. The presumed digital evidence scientific robustness and ability to corroborate every aspect of the investigation conceals limited accountability in digital forensics and undermines classical fair trial procedural guarantees in technologically complex and volumized evidence processing at an early stage of the investigation.

[2] Working definition based on legal and digital forensics perspectives elaborated from [Mifsud Bonnici et al., 2018, pp. 189–190], (I. ISO 2012), [Carrier, 2004, Pt. 2.1].

2. Inappropriate use of investigative technology

The Presumption of Innocence in technology-assisted investigations is firstly challenged given the inappropriate and inconsistent use of technology. This creates issues, some of which are already examined in (Findley and Scott, 2006; Risinger, 2008) such as tunnel vision[3] at an early stage of the investigation; lack of reliable and complete evidence; parallel construction of facts (Ciraco, 2001); lack of access to relevant evidence and forensic resources by the defence; unduly long retention of evidence and data on acquitted/suspected people for comparison. The need for harmonization of minimum procedural guarantees in respect to intrusive investigative measures is discussed at length in (Kusak, 2017; Vermeulen et al., 2010). Here the debate is enriched with considerations on the need for transparency and accountability in digital investigations and applied reliability validation of the digital forensic techniques employed.

2.1. *Datafication*

In the context of digital evidence, the PI has a role in strengthening the evidence at the investigation stage, where data is heavily processed and corroborated in growing amounts. Particularly problematic to PI are cases where the suspect or accused is suffering a limitation of his liberty or privacy based on vague suspicions. Besides side-stepping and erosion threats to PI, such measures can have adverse effects on the quality of the evidence and the trial in general.

The greater complexity of digital investigation and technology provides extensive access to and collection of "potential evidence", which is not examined by a court, but has a significant impact on the suspects, third parties, and even on the decision on whether, if at all, the cases will reach trial. For example, computer surveillance is the most intrusive investigation measure, because it interferes with the rights to privacy, data protection, and telecommunication secrecy, and in addition interrupts the integrity and confidentiality of computer systems. *Sunde* calls for regulation of such actions and refers to their difference from known interception or searches – "computer surveillance makes it possible to capture data that is not even intended to be transmitted (and could not be intercepted) and has not yet been stored (volatile data such as passwords and encryption keys) (and could not be seized)" [(Årnes, 2018),

[3] Tunnel vision is a "compendium of common heuristics and logical fallacies," that lead actors in the criminal justice system to "focus on a suspect, select and filter the evidence that will 'build a case' for conviction, while ignoring or suppressing evidence that points away from guilt". Definition in Dianne L. Martin, Lessons About Justice from the "Laboratory" of Wrongful Convictions: Tunnel Vision, the Construction of Guilt and Informer Evidence, 70 UMKC L. REV. 847, 848 (2002).

Ch. 3]. Moreover, the term surveillance is collectively used for all types of technology from police hacking to CCTV cameras, and even to data retention and monitoring.

To illustrate datafication effects in relation to PI, further examined are examples of *de facto* reverse burden of proof, data retention, and surplus of information.

2.1.1. *De facto reversed burden of proof*

ECtHR has stated that the reversal of the burden of proof is in principle forbidden;[4] however presumptions of fact and law, or asymmetric rules of proof as part of every legal system, could in some circumstances be unfavourable for the accused[5]. In *Murray* the court underlines that presumptions unfavourable to the accused are acceptable when the prosecution has established a strong *prima facie* case and when, according to the evidence, this is the only common-sense inference.[6] On the contrary, in *Telfner* the attempt by the prosecution to cover a lack of evidence by establishing presumption was sanctioned as a reverse burden of proof which violates the presumption of innocence. The judgements show that the Court is attentive to evidence irregularities in the investigation procedures. However, under the fourth instance limitation the court only states that legal presumptions depend on "the importance of what is at stake".[7] It fails to establish to what extent the presumption of innocence could be infringed in order to achieve other important goals in the criminal process, and when such infringements amount to violation. In digital context, this could be falsely interpreted as entailing the complete erosion of the PI in the name of security-focused and intrusive investigative measures designed to overcome technological barriers for prosecution. The use of technology allows to circumvent the prohibition of a reversed burden of proof by extensive use of probabilities and assumptions about "digital facts", where the reliability of digital evidence, its origin, or how it was obtained is challenged on legal, and not on forensic science grounds.

Although it is logical for reversed burden of proof to be used only as an exception and in minor cases [(Trechsel and Summers, 2006), Ch. 7], a report on recent evidence gathering practices shows that most countries are "lowering the thresholds (reasonable suspicion or serious indications to simple indications, reversed burden of proof, legal presumptions of guilt) for triggering the criminal investigation and for imposing coercive measures, the presumption *innocentiae* is undermined and replaced by objective security measures" (Vervaele, 2009).

Milaj and Mifsud Bonnici (2014) examine the use of several technologies to surveil and collect intelligence about targeted suspects and conclude that this undermines the PI principle and results in a *de facto* reverse burden of proof because of the danger of parallel construction of facts, collection of extensive personal information which undermines the right to remain silent, circumventing protective mechanisms in the criminal process, and "precooking" evidential material long before any charges are pressed. Some forms of criminal profiling may even result in a *de facto* presumption of guilt (Hildebrandt, 2014). The lack of access to information by the suspect to what is considered relevant in such "data expeditions" might prejudice any further adequate defence and denies any protection to individuals with unconventional behaviour who are not criminals. Moreover, a data-driven presumption of guilt has adverse effects on fairness because it does not situate the suspect as a party in the evidence-gathering process and requires the suspect to prove his innocence. It is also more difficult for the suspect to provide an explanation, cross-examine or gather counterevidence, since the evidence is "prepared" before any criminal proceedings take place. It could even be argued, that emotionally vulnerable suspects could suffer psychological damage and end up making false confessions.

Stuckenberg further argues that, in practice, given societal sensitivity and media pressure on the judge, the police and the state prosecution, insubstantial and questionable evidence is used to secure a conviction in an almost hysterical manner (as for example in the context of terrorist activity), or where the defendant is the only person to give evidence, this will result in a *de facto* reverse burden of proof to the disadvantage of the accused [(Stuckenberg, 1998), Ch. 3]. As *Gross* argues the "miscarriage of justice" occurs not on trial, but much earlier in the investigation. Time and social pressure can result in law enforcement striving for conviction and identifying the wrong person as the criminal. The amount of data available makes it easier to "gather enough evidence against this innocent suspect [and] the error will ripen into a criminal charge" (Gross, 1996). The impact of such misidentification is emphasised by tech-assisted investigations, where the line between preventive, security and investigation techniques is blurred.

Consequently, the investigation stage becomes longer and more complex, while a judicial warrant may not reflect the multiple dangers to the integrity of the investigation or the quality of the evidence. When the judge is not aware of the risks of certain technologies or investigation methodologies, is not informed why a specific method is preferred over others in a particular case, or about its accuracy, the judge is not in a position to evaluate the intrusiveness of the forensic technique and the warrant turns into a "blunt sword" of administrative compliance. The suspect could be prosecuted very differently depending on the law enforcement agencies' (LEA) digital forensics capabilities, lacks the protection of a formally-charged person, and is easily put in a position to have to prove her innocence.

Furthermore, the literature on conceptualizing reverse burdens is trial-based and does not resolve *de facto* reverse fact-finding during investigation. However, this does not mean that such conceptualization is not applicable to digital evidence investigations. For example, *Hamer* argues that when "the cost or probability of wrongful conviction is relatively low, and the cost or probability of mistaken acquittal relatively high, it may be necessary to lower the standard of proof, or even to reverse the burden of proof" in order to protect the PI. In digital and investigative perspective, such a principle can be transposed

[4] Philips v. The United Kingdom, Appl. no. 41087/98, §32, 5 July 2001; John Murray v. the United Kingdom [GC], Appl. No 18731/91, § 54, 8 February 1996, Reports 1996-I; and Telfner v. Austria, no. 33501/96, § 15, 20 March 2001.

[5] Salabiaku v. France, 7 October 1988, Series A no. 141-A, § 28: "Article 6 § 2 does not therefore regard presumptions of fact or of law provided for in the criminal law with indifference. It requires States to confine them within reasonable limits which take into account the importance of what is at stake and maintain the rights of the defense."

[6] John Murray v. the United Kingdom § 52 and § 60–62.

[7] Salabiaku v. France, § 28–29.

in proportionality assessment, as to whether the probability of reliable evidence discovery is high and the detriment to defendant/suspect rights infringement is low, that the asymmetric or reversal rule could be justified. However, further problems occur when the *de facto* reversal could infringe the rights of groups of people (targeted as suspects) or be at odds with the privilege against self-incrimination.

Further, the presumption of innocence has practical applications during the investigation in its "disciplinary effect in relation to the evaluation of evidence," and "verification of information" from different sources (Sliedregt, 2009). Given the identified complexities at the early stage or the pre-investigation use of technology, it is not a valid argument that law enforcement can use data processing just to assist the investigation or for intelligence, while fairness standards are triggered later with respect to classical investigation measures. Any use of technology during the investigation must meet a reliability standard which should be weighed against the probability of wrongful prosecution and error. This is primarily related to the rules on how to present and evaluate the reliability of evidence. *Jackson and Summers* argued that "where there has been a failure by the prosecution to obtain significant evidence or undertake various tests to establish the accused's guilt, the burden ought to be placed on the prosecution to prove why that has not prejudiced the defence" [(Jackson and Summers, 2012), Ch. 11]. Ergo, testing the evidence reliability by the prosecution is a way to avoid reverse burden of proof, while supporting digital evidence with reliability testing information serves the judge to further decide on the probative value and admissibility of such evidence. However, in the discussion of digital evidence standards, the evaluation of digital forensics best practices for their compliance with fair trial standards is not included (European Commission 2017).

2.1.2. *Data retention and surplus information*

The need for data retention for investigation purposes is well recognized by law enforcement authorities, but fundamentally questioned and criticised within the data protection community, which also affects cooperation between LEAs and the private sector. The controversial nature of data retention laws is partially rooted in the apparent inability of the legislator to guarantee sufficient safeguards, and maintain an appropriate necessity and proportionality test for data retention, which was also emphasised by the CJEU when invalidating the Data Retention Directive.[8] However, the UN report concluded that "national legal obligations and private sector data retention and disclosure polices vary widely by country, industry and type of data. Some countries report challenges in obtaining data from service providers." (United Nations Office on Drugs and Crime (UNODC) 2013).

In the Tele2Sverige case[9] the CJEU decided that a general obligation for collection of traffic and location data by all service providers for the purpose of combating crime is not in compliance with EU data protection legislation and required the collection to be limited to only what is strictly necessary and proportionate.

The classical portrayal of data retention practices by police as a privacy issue must be enriched by consideration of its impact on the PI. Firstly, there is the need to examine a couple of important safeguards in data retention practices formulated by the ECtHR and relevant to the PI discussion. The storage of data has to be subject to strict time limits even when it concerns serious crimes and individuals must have the opportunity to challenge the retention and the truthfulness of the records.[10] Moreover, the court underlined that the mere storing of data amounts to interference with Art.8 but failed to clarify the question of "subsequent use of stored data".[11] For example, questionable practices were described as a "function creep" or "surplus information"[12] where digital evidence collected for a certain purpose may end up being used for a different purpose. In Sweden, Finland and Denmark information collected during wire-tapping or computer surveillance, which exceeds the scope of the investigation, is not regulated by law. This surplus of information could be used as evidence in another case or serve for investigation and crime prevention purposes. The Swedish Council on Legislation (Lagrådet) has pointed out that specific regulation on the use of surplus information is needed in order to comply with the obligations under Art.8 ECHR.[13]

Further, in *Marper* the ECtHR dismissed the argument of the prosecution that specific technology and expert knowledge were not at their disposal to render the information intelligible – and concluded that the fact that the existence of such a possibility is sufficient to consider it as interference with Art.8.[14] Moreover, any data which is irrelevant for the purpose for which it is obtained must be immediately destroyed, storage of evidence data after the trial must be regulated by law and judicial authorisation is considered the main safeguard against arbitrary and abusive surveillance practices.[15] In addition, the Court endorsed the need of secure storage and security clearance for dissemination of interception material to be guaranteed.[16] The ECtHR underlined that even "*public information* can fall within the scope of private life where it is systematically collected and stored in files held by the authorities. That is also where such information concerns a person's

[8] CJEU, Judgment of the Court of 8 April 2014 in Joined Cases C-293/12 and C-594/12 Digital Rights Ireland and Seitlinger and Others, ECLI:EU:C:2014:238.

[9] CJEU, Judgment of the Court from 21 December 2016 in Joined Cases C-203/15 Tele2 Sverige AB v Postochtelestyrelsen and C-698/15 Secretary of State for the Home Department v Tom Watson and Others, § 108-109.

[10] S. and Marper v. The United Kingdom, no. 30562/04 and 30566/04, 4 December 2008; and Rotaru v. Romania (GC), no. 28341/95, 4 May 2000, § 43-44.

[11] Ibid., Marper v. the UK, § 67; and Amann v. Switzerland, no. 27798/95, 16 February 2000, § 69.

[12] FP7-SECT-2007-217862, DETECTER project, The use of surplus information in the court of law, 2007.

[13] EU Network of Independent Experts on Fundamental Rights, Opinion on the status of illegally obtained evidence in criminal procedures in the Member States of the European Union, CFR-CDF.opinion3-2003, available at:https://sites.uclouvain.be/cridho/documents/Avis.CFR-CDF/Avis2003/CFR-CDF.opinion3-2003.pdf

[14] Marper v. The UK, § 75.

[15] Roman Zakharov v. Russia, no. 47143/06, 4 December 2015, § 255-256.

[16] Kennedy v. The United Kingdom, no. 26839/05, 18 May 2010, § 162-163.

distant past."[17] The Court also outlined the particular danger of data collection with "the aim of being permanently kept and regularly processed by automated means for criminal-identification purposes."[18]

Currently, the impact on PI of data retention practices is not examined by the court but applying by analogy the logic in relation to Art.8 infringements, a couple of threats to PI are outlined here. One particular issue is that the systematic collection and retention of data could be easily misused for targeting suspects, reversing the burden of proof by requiring from the individual to prove she is not guilty and the data is false, for parallel construction, extortion of an innocent defendant, or as leverage when the prosecution lacks evidence. Secondly, the traditional safeguards such as the right to silence, the privilege against self-incrimination and even investigation procedures such as interrogation will be rendered useless and substituted by data analysis, where the data subject is rarely involved to exercise procedural rights as in criminal procedures. Thirdly, the issue with retaining data of different sources and origin, if they are not evaluated for their accuracy and reliability, has the potential to erode any protection of innocent people. The automated analysis of inaccurate data adds additional risks of errors and bias. The technical facilitation of data retention might also have significant impact on accuracy and result in tampering or data loss during storage, migration, or updates. This might also provoke the opposite effect where individuals guilty of crime are not identified due to low data quality and analysis.

Further, a lot of the work related to PI infringements facilitated by technology is focused on examining extreme cases such as mass surveillance or antiterrorist measures. As *de Hert* argues in the past "enhancing both the reliability and the 'softness' of surveillance measures contributes to their legal receptiveness and apparently silences civil liberty arguments" (De Hert, 2005). After examining data retention or the use of biometrics for security purposes as a form of "soft" surveillance, the author outlines the difficulties in applying the principles of proportionality and subsidiarity as legal tests for the intrusiveness of the measure. *Milaj and Bonnici* also argue that data retention laws are "*de facto* legitimizing a form of mass surveillance of citizens" (Milaj and Mifsud Bonnici, 2014). Such an opinion, however, does not account for the fact that more than half of internet traffic is encrypted. Improvements in anonymisation and obfuscation techniques greatly benefit criminals and are broadly used in dark web markets. Simultaneously, criminals and criminal behaviour develops and changes over time, while research in detection and identification of such behaviour is improving. Retention of data is crucial for forensic research, statistics on law enforcement technology use, and validation of digital forensic methods.

In summary, to face new threats to society facilitated by technology, states introduce both at the substantive and procedural levels tech-facilitated measures which lead to sidestepping and evasion threats to the presumption of innocence. In addition to data protection impact assessments, the presumption of innocence must be taken into consideration in further development of robust data retention legislation for law enforcement, specifically focused on the further processing of such data by automated means, the protection of the newly-inferred data, and the issue of repurposing and merging of data from different data bases and sources with different levels of accuracy.

In relation to the identified datafication effects, if the technology as such is not in breach of legal standards, the manner of its use may raise concerns.[19] Therefore, the use of technology is another factor to measure against fairness standards. The fairness of the investigation depends on the quality of the data. The great potential of more complex and volumized data for investigation purposes, could only be realised if a minimum level of data accuracy and reliability is assured through the presumption of innocence *by default* mechanisms, irrespective of data origin or technology, to counterbalance the impersonality and adverse effects of data in investigations. This requires a minimum standard for accountability and procedural accuracy for data processing which allows to trace back processing operations to systematically correct errors and verify inferences from different data sources.

2.2. *Technological circumvention*

An extensive and systematic capturing of data on suspects by the state renders other measures – which are well suited to the criminal procedure such as interrogation, witness examination or even detention – useless, since the state could "hack" its way into the most intimate details of suspects' lives. Preferring data instead of classical policing methods not only increases the danger of erroneous and premature conclusions in the investigation, but it avoids the criminal procedure altogether, because it diminishes the rights associated with an interrogation or witness statement rights and safeguards. Lack of "early and complete discovery disfavours the factually innocent far more than any other set of defendants" (Risinger and Risinger, 2021), as argued by *Risinger*, and this is particularly problematic in the current over-reliance on technology.

The use of technology may conceal the intrusiveness and significance of the LEA operation or cross-border cooperation. Under "technological means" the police could avoid limitations imposed by fairness standards or use transnational cooperation to circumvent domestic constitutional limitations. Cyber or non-physical aspects of the action may make LEAs investigative measure look less substantial than they really are, thereby pushing searches once considered joint ventures into the purely physical realm into primarily foreign searches given special status under the international "silver platter doctrine" (Street, 2011).

Hadjimatheou (2017) makes the important argument that not all technology-enabled interferences with individual rights invoke presumption of innocence safeguards. However, it is not considered that technology allows the combining of several non-intrusive measures, which can result in a high interference with individual rights. Indeed, Art.8–11 ECHR set

[17] Rotaru v. Romania, § 43 - 46. Emphasises mine.
[18] Marper v. The UK - in relation to fingerprint.
[19] Data Protection Working Party (DPWP) Archive, Opinion 1/2007 on the Green Paper on Detection Technologies in the Work of Law Enforcement, Customs and other Security Authorities.

important limits upon specific infringements during investigations. However, protection against accumulated irregularities due to an excessive use of technology must be covered by the PI, because it *de facto* amounts to the degree of criminalization and often towards whole groups in society. There is a need for oversight of the legality of a mounting up of IT investigation measures in time (periodically) or in analytical perspective (combination of personal, geospatial or other data with computation analytics, data mining). As argued by *Gless* a legislative failure is that "rules on information gathering only cover one law enforcement tool at a time, but never the whole picture that might, for instance, include a combination of telephone tapping, financial data mining, GPS surveillance etc." (Gless, 2010). Extensive use of technology can be used for PI evasion and side-stepping, and can potentially circumvent the criminal procedure all together, if fair trial safeguards and reliability standards are not implemented in the digital forensics processes. Intrusive digital forensic technology must be evaluated firstly atomistically – at each stage of the processing for fair trial compliance. More importantly, holistic evaluation about a periodical or analytical combination of measures must not be excessive in respect to PI.

2.3. *Evidence forum-shopping*

Data synced on multiple devices, backups, and cloud storage allow the same data to be retrieved from different providers and jurisdictions. If one country prohibits certain intrusive investigative measure, LEAs can use mutual assistance or mutual recognition instruments in order to acquire evidence from a country where such measure is lawful. Because countries apply a different standard for foreign evidence (Gless, 2013) and are not in a position to scrutinize a complex digital evidence processing in another jurisdiction, the legality or reliability of such evidence can be only challenged on limited grounds. Moreover, LEAs could decide not to intervene to stop a criminal activity in order to secure more convictions or a more severe conviction (Rumold, 2016). For example, the French LEAs used a computer interception device to break into an encrypted communication service Encrochat, allegedly used for facilitating criminal activities[20]. As a result, over 30 000 users[21] were intercepted in a period of 2-months and the data collection was further send to multiple jurisdictions in Europe which led to thousands of arrests in UK, The Netherlands, Sweden, and Norway (Zagaris and Plachta, 2020). However, the reliability of the interception evidence cannot be scrutinized since it is protected by military secret exemption. Although the investigation measure was authorised by two judicial orders in France[22], it is unclear if the authorisation encompasses the consequent broad computer surveillance. Moreover, the operation included data acquisition, analysis, and further processing among several stakeholders, including Europol. It is unclear who had access to the originally acquired data, how it was further analysed and filtered, and how it was attributed to concrete suspects. It is up to defence lawyers to challenge the evidence on a case-by-case bases which raises the question if the legality and reliability of the whole interception will be challenged at all. The Encrochat operation is an example of how little consideration is given to digital forensics reliability standards and defendants' right protection in technology-assisted cross-border investigations and mutual-trust cooperation mechanisms.

Defence lawyers have to deal with the threat of evidence forum-shopping because the prosecution services are embedded in formal trans-border networks which help them to find the best place to prosecute a case [(Boister, 2018), Para. 17.10]. So far, the PI as a principle has been examined to derive evidence rules at different stages and with respect to certain specifics of the investigation within the criminal procedure as a whole. The viability of such analysis must be tested in the context of cross-border investigations and cooperation for evidence collection and exchange, which is the standard scenario in the digital domain.

2.4. *Broadening investigative powers*

The increased use of encryption and emerging technologies (e.g. cloud) by criminals and third parties related to an investigation, creates the need for broader LEA powers such as anti-encryption laws[23] and data access laws.[24] The legality of such intrusive investigative measures is evaluated according to proportionality and necessity tests. However, it is questionable if such test can be applied when crucial information about the technology used or essential parts of the data processing are not disclosed or scrutinized in detail. Broad data access and decryption powers are not legislatively complemented with requirements for reliability and accuracy of digital forensics decryption methods or defendants' rights safeguards. For example, in relation to lawful hacking a suspect or defendant must be provided with information about the decryption methods and tools, about how the acquired decrypted data was examined and analysed, and potentially to be provided with DF assistance for testing exculpatory hypothesis against the data set. There is no standardized or approved methodology for forensic decryption, and in rela-

[20] Joint Eurojust-Europol press release, 'Dismantling of an encrypted network sends shockwaves through organised crime groups across Europe', 2 July 2020, https://www.eurojust.europa.eu/dismantling-encrypted-network-sends-shockwaves-through-organised-crime-groups-across-europe. Visited 2021-05-05.

[21] The exact numbers of intercepted messages and EncroChat users are unknown.

[22] Higher Regional Court Hamburg (Oberlandesgericht), decision of January 29, 2021, 1 Ws 2/21, 1 Ws 2/21 - 7 OBL 3/21, available at: http://www.landesrecht-hamburg.de/jportal/portal/page/bsharprod.psml?showdoccase=1&doc.id=JURE210003021&st=ent

[23] For example, Investigatory Powers Act 2016 of the United Kingdom; Telecommunications and Other Legislation Amendment (Assistance and Access) Act 2018 of Australia (The Parliament of Australia, 2018).

[24] Proposal for a Regulation of the European Parliament and of the Council on European production and preservation orders for electronic evidence in criminal matters COM(2018) 225 final, Strasbourg, 17.4.2018; and Proposal for a Directive of the European Parliament and of the Council laying down harmonised rules on the appointment of legal representatives for the purpose of gathering evidence in criminal proceedings, COM(2018) 226 final, Strasbourg, 17.4.2018.; In USA CLOUD Act, H.R. 4943, 115th Cong. div. V (2018) (enacted).

tion to mobile forensics exploiting vulnerabilities is currently the broadly used method by LEAs for access to mobile data (Alendal et al., 2018; Hay, 2017). Legislative initiatives are related to lawful hacking powers which are highly controversial in respect to human rights, intellectual property, security of communication (Koops and Kosta, 2018) and have international effects on digital governance policy (Budish et al., 2018).

Currently there is an over-emphasis both by scholars and legislators on access and collection of data through technology, while the debate is not legislatively complemented with reliability standard for testing of such data and its further analysis as evidence related to a concrete suspect or crime. Further steps in data processing after collection, specifically the legal compliance of pre-processing, examination and analysis of data (Broeders et al., 2017) from different sources in a law-enforcement context is not addressed in any legislative initiative. It is also unclear how information inferred from data (analytics) can be protected and used during the investigation.

2.5. *Challenging proportionality with reliability*

From the analysis so far, it can be seen that the use of technology could assist every step of the investigation process, but it blurs the lines between prevention and investigation of crimes and it might infringe the rights of suspects, groups of people or society as a whole.

Data protection laws increased the strictness of the principle of proportionality in respect to digital investigation measures: "collection of data on "contacts and associates" (i.e. on persons not suspected of involvement in a specific crime or of posing a threat), the collection of information through intrusive, secret means (telephone tapping and e-mail interception), and the use of "profiling" techniques, and indeed "preventive" policing generally, must be subject to a particularly strict "necessity" and "proportionality" test (Brown and Korff, 2009). In the latest communication, the Commission points out that the necessity and proportionality test depend on the type and volume of evidence, the type of investigation measure, its intrusiveness and safeguards of human rights.[25] It is not realistic, though, to assume that forensic experts can limit their collection methods only to relevant data from the beginning, or that all experts, who are often not lawyers, can successfully conduct a complex legal test. Simultaneously, cases where judges will need to refuse to accept important evidence due to the warrant's limits being exceeded, will need to be avoided. A clearer standard need to be developed to overcome this legal gap. As argued by *Hong and Yu* a "model needs to be established that can assess and regulate excessive search and seizure of digital evidence in accordance with a reasonable standard that considers practical limitations." (Hong et al., 2013).

The PI could greatly benefit from technologies designed to mitigate errors and uncertainties in digital evidence reconstruction, but it is often confronted with fairness principles such as proportionality, necessity, and subsidiarity. For example, in child pornography cases certain techniques focus on identifying the victim and how the illegal material is processed and stored by the suspect through automated comparison and searches in data bases (Lillis et al., 2018), while others focus on skin detection or examining messaging platforms to detect grooming chats and child pornography context (Amuchi et al. 2012; Ulges and Stahl, 2011). The issue is that we do not have sufficient documentation of the use of one or other technique in order to evaluate which one is more effective, less intrusive for the human rights of the suspects/victims or the rights of individuals at large. It is noticeable that analysis techniques become more reliable with a larger data set, which perversely is more privacy intrusive. One of the effects of data analytics is said to be that moderately effective algorithms produce better results from very large amounts of data than better algorithms do from smaller amounts of data. For example, advanced statistics give promising results in authorship attribution and detecting cyber bullies or predators online (Amuchi et al., 2012), but require increased monitoring of social messages and chats and a significant amount of sensitive data for training the algorithms. For example, studies show that the accuracy of feature selection or classification methods in text analysis depends on context, length of text, and number of authors (Layton et al., 2010; Zheng et al., 2006).

According to the *no free lunch theorem* there is no superior algorithm for solving a search problem if "their performance is averaged across all possible problems" (Igel, 2014). However, in digital forensics same tools are reused and repurposed for terrorism prevention, for child pornography detection, for investigation of murders, for example sentiment analysis, link and text analysis etc. This raises concerns about the accuracy, optimization, and validation of the implemented algorithms. Optimization of algorithm performance is based on understanding *(i)* under which conditions the algorithm achieves optimal results and *(ii)* the characteristics of the dataset. For example, some algorithms work better on linear and other on non-linear feature-feature or feature-class correlations (Nguyen et al., 2010). If an algorithm was originally designed under the assumption of a linear correlation yet the dataset has non-linear correlation the results will be suboptimal. On the second precondition, the accuracy of the search algorithm depends on prior "problem-specific knowledge to achieve better than random performance" (Wolpert and Macready, 2005). This means that the accuracy of examined digital forensic problem will depend on the algorithm selection and correct interpretation and assumptions about the input data based on knowledge specific for e.g. child pornography, murder, or terrorism cases. All-purpose forensic tools are not designed for domain specific investigation, does not make the assumptions about the data set characteristics explicit, and have a fixed algorithm implementation. Moreover, if they are closed source there is no possibility to validate the process of algorithm and feature selection except if the source code is not disclosed. Furthermore, without the information necessary for technical validation of methods and tools, the legal assessment of legality, proportionality or reliability of data-hungry algorithms is not possible. This indicates a level of testing and reliability eval-

[25] Council of the European Union, 9554/17 Technical document: Measures to improve cross-border access to electronic evidence for criminal investigations following the adoption of the Council Conclusions on Improving Criminal Justice in Cyberspace, Brussels, 22 May 2017, part V.C., p.46.

uation of methods in computations combined with human rights impact assessment which as observed further is not yet achieved in digital forensics.

Little is known about which technologies in digital investigations are excessive in relation to the PI. The notion of a "reasonable suspicion" is insufficient and imprecise to satisfy forensic data analysis. The proportionality principle and balancing tests are notoriously vague and risk becoming completely impractical for scrutinizing digital forensics processing. In the absence of a legislative approach, such a complex legal evaluation is transgressed to law enforcement, police or digital forensic specialists, which is contrary to the presumption of innocence and the rule of law.

Any new digital forensics technology for law enforcement purposes must be scrutinized from its design to its application for accuracy, intrusiveness, and PI compliance, which requires a transition from "all-purpose tools" to "*digital forensic methods for law enforcement*". In a legal analysis, it must be considered that the use of technology in an investigation could interfere with or violate the PI. This requires assessing when the interference is justified or not, what error-mitigation mechanisms exist on both pre-trial and trial to counter pervasive investigation techniques and compare the reliability versus the intrusiveness of the digital evidence processing.

3. Old procedural guarantees vs. new digital evidence processes

A set of PI threats is invoked by old procedural guarantees on trial, which are not adapted to contemporary digital evidence processes and services.

The ECtHR has always emphasised the importance of the principle of orality, where independently of the statements or discoveries in the investigation, on trial the "evidence must be produced [...again] in the presence of the accused at a public hearing with a view to adversarial argument."[26] A further requirement is that "the whole matter of the taking and presentation of evidence must be looked at in the light of paragraphs 2 and 3 of Article 6."[27] The burden of proof requires also the prosecution to take "positive steps" for disclosure of evidence and respect the defendant's procedural rights. Moreover, in *Barberà* the Court objected that "1600 pages investigation file, the bulk of which did not concern the defendants"[28] does not meet the disclosure requirement. What is required are the personalized findings of the prosecutor, to "specify in detail the particular evidence on which he based his account of the facts in relation to the defendants," as well as disclosure of exculpatory evidence and unused material. These requirements and the effectiveness of the principle of orality and the right to challenge evidence on trial in the digital age are shaken and as argued further ineffective for digital evidence processes.

In a recent case, the prosecutor argued that due to the volume of data and problems with reimaging the data, the defendant should be provided only with the data selected for the examiner's files.[29] Moreover, the investigators conducted several searches with forensic tools, without judicial supervision or involvement of the defence, where the defence had no possibility to replicate the prosecution searches or search the whole data set for exculpatory evidence. The Court underlined that the prosecution arguments were not justified and that all this "would in principle raise an issue under Article 6 § 3(b)". However, in the particular case no violation of Article 6 was found given the inadequate way of challenging the digital evidence[30] – the defence failed to obtain a court order to access the full data set or had not suggested further investigative measures – such as a fresh search using keywords suggested by them. The two cases exemplify that when dealing with data sets and forensic examination if the defence does not get a chance to cross-examine the procedure for deriving evidence from data and search for exculpatory evidence on pre-trial, the opportunity to find data as evidence or to scrutinize prosecution actions will be lost. The Court's emphasis on the importance of a procedure to challenge digital evidence on pre-trial is a sign of the erosion of the orality trial guarantee due to the volumes and processing of evidence data in investigations, and an instruction to strengthen the active defence evidence rights on pre-trial stage.

The trial guarantees such as the principle of orality, disclosure, and cross-examination of digital evidence are equipped to examine only the results of digital forensic processes in relation to the legal arguments of the case. It is questionable if the requirement for "reasoned judgements" can validate the factual accuracy of the digital evidence, and if the legal understanding of the judge of the forensic report summary is sufficient to take an informed decision considering that the digital forensic process is largely omitted from the evidence report. Furthermore, defence lawyers are also prevented from challenging digital evidence on reasonable grounds in case they had no information on the data sets, tools and methods used but rely only on the reported results and the expert opinion. Foster and Huber argue that "cross-examination, presentation of contrary evidence, and careful instruction on the burden of proof are the traditional and appropriate means of attacking shaky but admissible evidence ... This language is usually cited by those favouring looser standards of admissibility" (Foster and Huber, 1999). By analogy, loose standards on the reliability of digital evidence and digital forensic process cannot be compensated by the defence rights on trial.

Weak procedural mechanisms to scrutinize digital evidence on trial or pre-trial often result in taking the relevance and probative weight of the expert testimony for granted, while challenging it on reasonable grounds requires a level of technical literacy (Edmond and Roberts, 2011). In most jurisdictions there are no clear or effective procedures to challenge expert reports or to examine the reliability of the scientific findings (Henseler and van Loenhout, 2018; Marsico, 2004; Vuille, 2013). Procedural guarantees such as the principle of

[26] Kostovski v The Netherlands, no. 11454/85, 20 November 1989; Barberà, Messegué and Jabardo v. Spain, no. 10590/83, 6 December 1988.

[27] Barberà v. Spain, § 78; also, Capeau v. Belgium, no. 42914/98, 13 January 2005, § 25.

[28] Barberà v. Spain, § 77.

[29] Sigurður Einarsson and Others v. Iceland, no 39757/15, 4 June 2019.

[30] Ibid., § 91-92.

orality, disclosure and cross-examination of digital evidence do not really work with respect to digital data which is heavily pre-processed long before trial. Digital forensics is a discipline which is constantly fighting against resources deficits, which require the implementation of verification and validation mechanisms for legal compliance to be implemented in evidence processes and systems. Despite admissibility differences between jurisdictions, the requirement for reliability of digital forensics turns into the main instrument for international discussion and harmonization.

Moreover, due to the digitalization of most critical services in society, digital evidence increasingly must be accessed, used, and understood by suspects, accused, and defendants. *Van Wijk* examines the equality of arms principle in the context of evidence gathering with an international component and concludes that as a safeguard for the accused she has to be able to either actively by gathering evidence by herself or at least passively, with the assistance of law enforcement authority, be able to collect evidence cross-border (van Wijk, 2017). However, this remedy could be limited when the defence has a heavy burden to prove the exact scope and location of the digital data, or the particular importance of the requested digital evidence to the case. Further, the use of technology may conceal the significance of the LEA operation, cooperation or intrusiveness. Due to increased cross-border cooperation in investigations, the regime of foreign evidence must be harmonized and improved with reliability standards for digital forensics. This will prevent the danger of parallel construction or the admissibility of illegally-obtained digital evidence from abroad (Gless, 2013).

Therefore, it is further considered that the employment of computational methods in investigations requires a level of accountability and transparency which can facilitate an evaluation of the proportionality and reliability of the method and ensure a better position for suspects/defendants and judges with respect to the technology used and its results in digital evidence. Reliability criteria documentation, although only a first step in such an analysis, is a crucial component for effective and fair tech-assisted prosecution. However, despite legislative and standardization requirements for reliability testing and "forensically sound" procedures, the practice shows deficits and a lack of solutions for validating tools, examiners, and methods, as well as a lack of sufficient documentation and standardized procedures for a reliability assessment.

4. Digital forensics –a threat to the PI?

Shapiro identifies that one of the greatest challenges in the historical understanding of facts for litigation was the transit "from something that had to be sufficiently proved by appropriate evidence to be considered worthy of belief to something for which appropriate verification had already taken place" [(Shapiro, 2003), p. 31]. Verification procedures for the reliability and authenticity of digital evidence became a topic broadly discussed for practical solutions that will ensure protection against wrongful convictions and low-probative evidence. Digital forensics is defined as "the use of scientifically derived and proven methods towards the preservation, collection, validation, identification, analysis, interpretation, documentation and presentation of digital evidence derived from digital sources." In forensics, reliability is a property of a process related to consistent intended behaviour and results. (I. ISO 2012) It equals reproducibility and validation of the method (Foster and Huber, 1999; Risinger, 2018). A test that produces the same results on successive applications is said to be *reliable* (Gross and Mnookin, 2003).

The evidence rationalists see the need to complement the trial-based evidence evaluation, with intellectual developments in adjacent forensic fields, standards for non-adjudicative decisions involving a fact-determination especially on pre-trial phase, *and the extra-judicial significance of the presumption of innocence* [(Twining, 2006), p. 243]. Ideally, the criminal process as a whole follows the PI principle, but also requires its safeguards implementation in a multidisciplinary context. This paper examined how *Ashworth's* four legal threats to PI are facilitated by technology, and this concept could be further elaborated in regard to the scientific threats to PI.

The presumption of innocence as a rule of treatment can be interpreted in digital context as a requirement to procedurally mitigate and minimize the impact of bias, uncertainties, and errors in digital forensic science. Since digital evidence is the result of machine-human interactions, standard procedures for digital forensics must mitigate both machine and human errors. To the contrary, judges "seem to be enthusiastic to rapidly embrace the products of technological progress" (Edmond and Roberts, 2011) and often assume that the digital media source of evidence is "working properly" [(Mason and Seng, 2017), Para. 6.198]. Similarly, "law enforcement and prosecuting authorities are often willing to use novel science and technology" in order to secure evidence [(Doyle, 2019), Ch. 7]. The enhanced use of automated tools to acquire and analyse digital evidence creates the false perception that technology mitigates errors and bias, and results from tools are reliable and trustworthy. In fact, such a *technological protection fallacy* (Dror, 2020) does not account for the multitude of errors, bias, and uncertainties in digital investigation that might have detrimental effect on the presumption of innocence and the fair trial in general.

Authors examine multiple biasing factors in digital investigations such as exposure to case-irrelevant information, base rate expectations from previous investigations, failure to evaluate competitive hypothesises, or digital context information indicating intent or bad character (Edmond, 2016; Sunde and Dror, 2019). Examiner errors are related to inaccurate data examination and tools result interpretation, as well as improper parameterisation of the tool. Tool and method errors are even more subtle and require formal validation to be detected. Since same tool or method for data examination can be used in multiple investigations and trials, failing to identify limitations and errors can potentially result in reopening all previous cases for re-examination once the errors are detected. Currently, there is no standard in DF for "calculating error rates for both tools and specific procedures" (Carrier, 2002). Moreover, errors related to many digital forensic activities are "systematic in nature and no statistical error rate exists" (Lyle, 2010). Therefore, cross-verification of results based on documenting the methods and tools, and the examiner interaction

in the process is often more reliable than calculating statistical error rate.

The fact that the reliability of digital data or digital forensics tools can never be assumed is examined further in detail to expose the insufficient scientific rigor and validation studies in the digital forensics discipline. The way digital forensic "science" is performed in practice shows a process led by operational and investigative objectives, which lacks scientific validity and process quality assurance.

4.1. *The erroneous concept of "digital forensic investigation"*

Traditionally, investigation is strictly separated from forensic examination – they are different career paths, requiring different expertise. The investigator is guided by law enforcement objectives, while forensic scientists apply scientific methods and aim at impartiality.

However, a consistent body of digital forensics literature departed from this legal and regulatory tradition by introducing the term "digital forensic investigation" (DFI) (Carrier, 2006; Ieong, 2006; Kohn, Eloff and Eloff, 2013; Montasari, 2016). Often the term "digital investigators" is used in the sense of digital forensic scientists (van Baar et al., 2014). DFI is defined as "the process to determine and relate extracted information and digital evidence to establish factual information for judicial review" (Ieong, 2006), that is identical to the traditional definition of forensic science. So why was the generic term "digital forensics (science) abandoned in favour of the new term DFI?

The need to introduce this new concept was described by Carrier as follows:

> *"Typical forensic science areas answer comparison questions. Unknown object is compared to a standard reference and the scientist determines if they are the same. An object is identified by comparing it to several references. The process that occurs in "digital forensics" on the other hand, involves searching for evidence, identifying it, and reconstructing events. The identification and comparison process is only one part of the big picture…"*

Based on this view, Kohn explained further that DFI is a "special type of investigation where the scientific procedures followed and techniques used will allow the results - digital evidence - to be admissible in a court of law" (Kohn et al., 2013).

Although, supported by numerous articles, these definitions are laying out an erroneous view of forensic science and more importantly mix together investigative and scientific objectives in criminal investigations, which prevents the validation and accountability of both. Firstly, forensic science is not dealing only with comparison questions. Legislators and standardization bodies have identified the core forensic science processes as: authentication, identification, classification, reconstruction, and evaluation of traces (Pollitt et al. 2018; Risinger, 2018). Secondly, the investigation objective is to present evidence admissible in court, but this is not an objective of digital forensics as a science. The objective of digital forensics is to test the strength of the digital artefact and events relevant to the investigative case. In fact, mixing the investigative and the digital forensics science objectives is described as a biasing factor (Sunde and Dror, 2019). Moreover, the artificial separation of the physical and digital investigation is confusing and does not reflect reality.

The investigation at a case level has different objectives and depends on the criminal procedure and the type of crime. By contrast, the digital forensics process model is suitable for standardization because it aims to ensure scientific validity irrespective of jurisdiction. It will also assist judges to evaluate the forensic methodology according to formalized procedure.

In practice, digital forensics is often conducted by law enforcement officers who are not digital forensic scientists (Adams et al., 2013; Montasari, 2016). Most of the forensics labs are either part of or financially dependent on law enforcement [(Doyle, 2019), Ch. 7]. Moreover, there are significant differences between the investigation objectives and digital forensics as a science. The fact that they are performed as one makes the quality evaluation of both harder, poses questions about professional bias, protection of innocent defendants and equality of arms in respect to digital forensics aid for the defence. Investigation is a process that develops and tests hypotheses to answer questions about events that occurred (Carrier, 2004). Investigators can search and observe digital traces and form hypotheses about the case. However, they lack competence to attribute, evaluate, interpret, and reconstruct digital traces (van Baar et al., 2014). Digital forensic scientists, on the other hand, evaluate traces (facts) in order to establish their probative strength (Pollitt et al., 2018). There is also an opposite effect of forensics scientists exceeding their domain expertise and becoming general investigators. (Risinger et al., 2002).

In conclusion, we advocate a clear separation between investigation and digital forensics as a science. The further focus in this paper is on challenges to digital forensics as a science producing expert evidence in court proceedings. However, the quality of the investigations in the digital domain and their upholding of fair trial standards, although not in the scope of this paper – requires further research, especially with respect to the EU efforts for harmonisation and cooperation in criminal procedures.

4.2. *Reliability crisis in digital forensics?*

In the legal domain several issues with unreliable forensic evidence are reported and discussed at length. Several reports have concluded that false confessions and unreliable forensic science evidence are factors in wrongful convictions (Edmond, 20162012; Innocence Project 2020). Moreover, academics argue about systematic over-estimation of the weight of expert evidence (Callen, 2015; Edmond, 2016; Edmond and Roberts, 2011). In most jurisdictions, judges continue to be provided with no real guidance on how they should determine evidential reliability (Horsman, 2018a) which also leads to unequal treatment of suspects and defendants (Gross and Mnookin, 2003; Risinger, 2000). Arguably, the outlined "classical" problems with all forensic sciences in adjudication are deepened in digital forensics given some specifics in digital forensics practice, not typical for other forensic disciplines.

Doyle (2019) conducted extensive research on the quality management of forensic science and its relation to fairness and concluded that the major challenges faced currently in all forensic fields are: the premature use of novel science and

technology which lies outside a quality standards framework, lack of standardization and harmonization, lack of resources, and accountability.

These issues were further emphasized by Interpol as serious challenges in the digital evidence domain (Reedy, 2020) and in the UK National Digital forensics Strategy [(T. F. The UK National Police Chiefs Council 2020), p. 21]. Digital forensics practitioners and academics expressed concerns about the lack of scientific validation in digital forensics (Casey, 2019; Horsman, 2018b; Hughes and Karabiyik, 2020; Jones and Vidalis, 2019), while the reproducibility crisis in the field was commented on by standardization and governmental bodies worldwide (Council of the European Union 2016; PCAST 2020). Several legal scholars called for digital forensics expert accreditation (Henseler and van Loenhout, 2018; Kwakman et al., 2011) and discussed the absence of clear legal rules for evidence reliability assessment to the disadvantage of suspects and defendants (Edmond, 2012; Risinger, 20182000; Saks and Koehler, 2005; Sommer, 2010). The rapid scientific advances in computer-assisted forensic science render a lot of existing validation schemas outdated (Kloosterman et al., 2015), side-track reproducibility studies (Horsman, 2019a; Tully et al., 2020), disturb accuracy testing in digital forensics (Garfinkel et al., 2009), and in the subsequent court evaluation (Saks and Faigman, 2008). The lack of resources to deal with these continues to grow data volumes, and complexity (Horsman 2018a; Jones and Vidalis, 2019), digital evidence dynamics, and data volatility (Horsman, 2019a; Morgan, 2017) is often used as an argument for not implementing quality standards (Horsman, 2019b 2018a; Casey, 2019).

Table 1 below summarises six classes of problems related to reliability assurance in the digital forensics' domain: procedure, tool, method, examiner, documentation, and data set. They all introduce accountability and transparency issues in respect to the investigation, and lack of compliance with scientific requirements for forensic evidence. Consequently, the DF in practice does not meet any fairness objectives; moreover, there is a lack any procedural assurance for the rights of the suspects and defendants.

The analysis shows that each of the classes is a suitable criterion for development of a formal validation matrix. However, the identified gaps show that in digital forensics testing prior to the use of a tool or method for law enforcement purposes is not sufficient to validate the results. Each of the validation classes must be tested during the forensic task on a case-by-case basis for its reliability. The specifics of the evidence data set or examiner interaction may impact the accuracy of the forensic methodology. These effects in the digital domain are symptomatic of the need to implement and enforce reliability testing in criminal investigation procedures. Minimum standards for procedural accuracy including strengthening the defendants' position in respect to digital forensics evidence is a necessary step in order to preserve fairness standards in technology-assisted investigations.

4.3. *Evidential reasoning*

Reasoning about digital evidence, especially when it is based on processing of data, is exposed to a high level of uncertainties and probabilistic inferences which needs to be examined for accuracy as well. Moreover, digital evidence is always the result of an interpretation either by the tool or by the examiner.

The similarities between criminal law and a programming language are quite interesting, and although criminal law could trigger a great variety of scenarios, it is the part of law which aims at maximum bivalent 0 or 1 answers, in order to ensure equal treatment of individuals and protect them from the random exercise of powers. A matrix that express the criminal law rule in standard language can be derived as follows:

> ***If <control parameters/> - [abstract legal rule]***
> ***match <input/> - [concrete parameters – facts] – uncertainty***
> ***then <output/> [consequences – punishment]***

The question is then how law, tools, and procedure deal with this uncertainty during the interpretation of digital artefacts. Shum argues that "the evidence is incomplete on matters relevant to our conclusions, and it comes to us from sources (including our own observations), that are, for various reasons, not completely credible. Thus, inference from such evidence can only be probabilistic in nature…" (Schum, 2001).

Stein argues that the mere probability of statistical evidence "irrespective of how high it is numerically, can never provide an adequate foundation for criminal convictions", therefore if essential evidence is missing or it is not "susceptible to maximal individualized testing" or "the testing undermines its credibility" exposes the defendant to an illegitimate risk of erroneous conviction. He further argues that allocation of "the risk of error in criminal … trials cannot derive from judicial forecasting of the success or failure of the ongoing scientific evolution or revolution" (Stein, 2005). However, the benefit of the doubt does not equal that someone is innocent or that someone is probably guilty. Such a probability analysis of the case could be taken into consideration when developing data retention rules or allowing enrichment of cold cases with new information.

Consequently, if the state cannot develop an adequate testing standard for evidence, before or during trial, this can result in an infringement or violation of the PI and fairness in general. Moreover, digital investigations are distant from the real world and rely on a generalization instead of individualized, trace-based investigation – they could be highly inaccurate if the presumption of innocence is not reinforced at an early stage of the investigation with respect to potential evidence and potential evidence sources. This raises the question of the reliability of the predictive (statistical) evidence versus the requirement of a trace-based fact-finding.

Jackson and Summers discuss rationalist theories about reasoning of innocence and guilt, which can be applied by analogy to the scientific endeavour in digital forensics when reasoning about the digital data itself (inference about digital evidence). In the examined probabilistic or falsification approaches "guilt, is measured by the relative amount of evidence for or against it *or* by subjecting the hypothesis to a variety of tests to determine its explanatory force in relation to the evidence" [(Jackson and Summers, 2012), p. 215]. In their view, the difficulty of accessing "how much" probability is required *or* how plausible the fact explanation should be to raise (rebut) reasonable doubt, are "shifting the fo-

Table 1 – Reliability challenges in Digital forensics.

Class	Challenges
Procedure	-need for standardization, field governance, entry requirements, and transparency (Horsman, 2018a; Doyle, 2019; Horsman, 2019b; Vincze, 2016) -increasing requirements to prove validity are burdensome for practitioners (Horsman, 2018a; Hughes and Karabiyik, 2020) -dissemination of unreliable knowledge (Horsman, 2019b2018a) -ISO 17,025 is the incorrect vehicle for regulation of standards in DF …but currently the only attempt to achieve a measure of quality control (Page et al., 2019; Jones and Vidalis, 2019) -lack of sufficiently large, centralized validation efforts and a lack of reproducibility studies (Hughes and Karabiyik, 2020; PCAST 2020; Page et al., 2019)
Tool	-underlying SW/FS remain unsupported by DF tools for a significant period of time (Horsman, 2019a) -validation procedures become rapidly obsolete due to versioning/ updates in all SW (Horsman, 2018a; Doyle, 2019) -limited versions of the same tool are tested (newer versions/ updates are not) (Horsman, 2019a) -commercial forensic tools are not tested for security vulnerabilities (Wundram et al., 2013) -testing is time and resource consuming (Horsman, 2019a) -tests are narrowly defined covering limited scenarios (Horsman, 2019a) -larger, commercial tools for verification can be both too costly and time consuming, whereas the smaller tools available can require too much interaction (Friheim, 2016) -reuse of libraries/ functions makes dual-tool verification obsolete - even where code can be accessed for analysis, it is likely that a practitioner would have neither the time nor resources to effectively scrutinise its structure for error validation (Horsman, 2018a) -full source code level audit of any tool, let alone any operating system component, to ensure precise and correct operation is basically impossible (Gerber and Leeson, 2004) the tool producers are either unable (in the case of most small providers) or unwilling (in the case of most larger providers) to provide information about how they capture customer requirements, let alone disclose what those requirements are (Marshall and Paige, 2018; Tully et al., 2020) -the investigator and the courts must trust that the digital forensic software/ hardware was created accurately (Marsico, 2004; Carrier, 2002; Patel and ó. Ciardhuaéin, 2000) -lack of proof and verification that a tool is treating all the input data in the same way, does not omit any data and processes everything according to the forensic objectives, and does not serve personal or corporate interests (Stoykova and Franke, 2020) -not all DF tools report errors or inaccuracies
Method	-need of sufficiently reliable scientific basis for the expert evidence (Horsman, 2018a) -no existing programme can demonstrate the foundational validity of digital forensic tools, nor provides an examiner or laboratory the resources needed to perform a comprehensive validation study of a particular implementation of a tool or method. (Hughes and Karabiyik, 2020) -limitations of repeatability (could repeat the same error due to code reuse) (Horsman, 2019a) -without knowing the algorithms that have been used in the tools, there is no way to ascertain that they are not using the same algorithm and are, in effect, self-validating (Jones and Vidalis, 2019) -at best the process incorporates a sampled-set of verified results (Horsman, 2019a) -lack of resource of ground truth data (Tully et al., 2020; Horsman, 2019a; Hughes and Karabiyik, 2020) -significant practical gaps exist in the ability to carry out tool validation in a scientifically defensible way (Hughes and Karabiyik, 2020) -examination of digital evidence as an investigative rather than forensic activity (Hughes and Karabiyik, 2020) -formal validation methodology, it may be wise to divide forensic tasks into sub-tasks (Hughes and Karabiyik, 2020) -inability to distinguish tool errors from user errors (Hughes and Karabiyik, 2020; Horsman, 2019a) -ideal condition validity vs. actual practice validity (Risinger, 2018) -challenge to know how to interpret the data (Horsman, 2019b; Sremack, 2007) -transparency and substantial disclosure (Edmond, 2016) -need of multidisciplinary peer-review (Edmond, 2016) -rapid method development and uncontrolled introduction of untested methods (Sremack, 2007; Arshad et al., 2018) - have not satisfied the criteria of known error rates (Jones and Vidalis, 2019)
Examiner	-quality systems which exist in DF often target organizations at a laboratory level, not the individual or their outputs and, as a result, it is difficult to ascertain the quality and reliability of investigatory work (Page et al., 2019) -'push-button' forensics – examiners' dependence on digital forensic tools (Horsman, 2019b2019a) -interaction with the DF tool and data set -expression of confidence and reporting error (Edmond et al., 2010) -DF practitioner may be responsible for all stages, where oversight and evaluation of the work carried out at each point may be minimal (Page et al., 2019) -the lack of minimum universal qualifications for digital forensic examiners, (Hughes and Karabiyik, 2020) -wide variation in education, experience, and training amongst digital forensic examiners (Hughes and Karabiyik, 2020) -examiner error and cognitive bias (Hughes and Karabiyik, 2020; Sunde and Dror, 2019)
Documentation	-time and resource consuming documentation overburdens practitioners (Horsman, 2019b2018; Casey, 2019) -poor documentation cannot serve as established practice (Horsman, 2018a)
Data Set	-accurate representation of the original data, its authenticity and integrity can cannot be validated (Casey, 2007) -minimum and identifiable modifications, inaccuracies, errors (Mocas, 2004; Beebe, 2009)

cus away from evidentiary tests for determining guilt towards examining the accuracy of the procedures that are adopted within the criminal justice system for determining guilt".

Thaman elaborates that "procedural rules should require a minimum amount of solid evidence of guilt in order to 'rebut' the presumption of innocence and send a case to a trier of fact to its subjective assessment of proof beyond reasonable doubt" [(Ross and Thaman, 2018), p. 76]. He even goes further in proposing "negative formal rules of evidence", which in capital crimes will prevent convictions on weak, circumstantial evidence, but he does not further specify criteria for the required sufficiency of evidence. While protecting the innocent defender from the risk of low-probative evidence can indeed be achieved by requiring standards and quality in the investigation procedure which allocates the risk of error to the prosecution, the need of strict exclusionary rules to enforce this is questionable, especially in a dynamic techno-legal domain like digital evidence. ECtHR case law repeatedly confirms that exclusionary rules on evidence do not form part of the European Public Order [(Jackson and Summers, 2012), p. 215]. Moreover, the introduction of PI-based evidence rules at the investigation stage benefits quality assurance and reliability testing, which reduce the need for exclusion. Further, throughout the analysis it has been established that the use of technology to assist investigations requires a scientific validation that can cope with the fast developments and changes in the domain and is easily verifiable by the court on trial, while strict exclusion will not account for the complexities in such a validation.

The interpretation of digital data by LEAs and DF tools for the purpose of the investigation must be made accountable and part of the validation procedure. Reasoning about the evidence during analysis and any assumptions made must be documented. Insufficient individualization and accuracy testing of the digital evidence can potentially result in fairness violations.

5. Current approaches to address threats to fairness

Further, the benefits and drawbacks of proposed legal and non-legal standards to ensure accuracy and fairness of procedure in different forensic disciplines should be outlined. In this context, the common understanding of the burden of proof as a mechanism to allocate the risk of error in the criminal process, is dependent on accountability and integrity policies for all tools, processes and services supporting the investigation.

5.1. *Reliability evaluation by the court*

Given the uncertainties about digital data, its reliability on the source or during forensic processing must be never presumed but proved and recorded. One proposed solution was for the court to do a special evaluation of reliability as in other cases when scientific evidence is in question.

The US Supreme Court formulated Daubert's rule,[31] which was the first decision to promote court criteria for evaluating expert evidence similar to those that scientists use. The rule influenced many countries internationally (Jasanoff, 2005), because it drew attention to wrongful convictions based on unreliable expert evidence (Innocence Project 2020). In recent years detailed reliability requirements were introduced in academia, forensics, and standardization bodies (Sommer, 2010). Common law countries focused on forensic evidence reliability requirements as a precondition for admissibility (Edmond, 2012; Edmond, 2011), while inquisitorial systems still rely on accreditation and certification of forensic experts (Kwakman et al., 2011).

This section examines the Daubert criteria in detail since it was exposed to serious criticism. This criticism is addressed here by arguing that Daubert was a step in the right direction, with high relevance for the digital evidence domain; however, the problem is the lack of standardized processes to generate information for scientific validation and the lack of formal validation procedures during investigations, which makes the work of the cross-examiner and judges impossible.

The Daubert standard requires: *(1)* the forensic theory or technique to be tested, *(2)* peer-reviewed, *(3)* generally accepted in the scientific community, *(4)* account for error rates *(5)* within the examiner's expertise. Numerous scholars expressed concern about the practical applicability of Daubert (Edmond, 2012; Risinger, 2000). The critics were mainly in three directions: *(i)* that the criteria in Daubert is unclear; *(ii)* that judges are ill-equipped to evaluate complex scientific methodology; and *(iii)* that Daubert seems to have made it more difficult for the defence to adduce expert evidence (Edmond and Roberts, 2011; Risinger, 2000; Jasanoff, 2005). As shown in Table 2 below application of this criteria in digital forensics outlines the limitations of procedures and a lack of standards to produce the information needed for a Daubert evaluation.

Digital forensics lacks the needed underlying scientific validation process in order to meet any of the criteria. Moreover, court proceedings are not equipped and judges have different objectives during trial than to deal with the complexity in such validation.

Daubert's principle is further questioned for its practical use in the absence of a "standard [...] established and certified by the justice system" on the required scientific evaluation to satisfy the court. Opinions were expressed that Daubert places judges as "amateur scientists" to evaluate complex scientific findings in checklist fashion (SKAPP 2003). Most state courts in the US have rejected Daubert and judges expressed the concerns that:

> *"Our responsibility then, unless we badly misread the Supreme Court's opinion [in Daubert], is to resolve disputes among respected, well credentialed scientists about matters squarely within their expertise, in areas where there is no scientific consensus what is and what is not "good science," and occasionally*

[31] Daubert v. Merrell Dow Pharmaceuticals, Inc., 509 U.S. 579, 1993. The Daubert criteria was further elaborated in General Electric Co. v. Joiner 522 U.S. 136 (1997), and Kumho Tire Co. v. Carmichael 526 U.S. 137 (1999).

Table 2 – Implementing Daubert in practice?

Testing	-lack of solutions to perform validation testing (Hughes and Karabiyik, 2020) -unclear amount, quality, quantity, or type of data needed for validation (Hughes and Karabiyik, 2020) -dual-tool verification is failing due to reuse of libraries/functionalities (Jones and Vidalis, 2019; Marshall and Paige, 2018) -test scenarios does not cover all functionalities of tools (Horsman, 2019a) -lack of time and resources results in lack of quality standards (Horsman, 2019b; Horsman, 2018a)
Peer-review	-what type of peer review is acceptable (Tully et al., 2020) -no common understanding of what makes the reviewer an expert (Marsico, 2004) -in DF the rate of change is faster than the time required for peer-review (Horsman, 2018a) - reference to peer-reviewed method in the DF report is insufficient, because it does not explain how the method was applied in the particular case. (Carrier, 2002)
General acceptance	-most unclear, even irrelevant requirement (Carrier, 2002) -proliferation of methods and practices in many areas of digital forensics, where none is considered a standard (Marsico, 2004; Horsman, 2019b; Sremack, 2007; Arshad et al., 2018) -no existing programme can demonstrate the foundational validity of digital forensic tools, nor provide an examiner or laboratory the resources needed to perform a comprehensive validation study of a particular implementation of a tool or method. (Hughes and Karabiyik, 2020)
Error rates	-lack of methodology to evaluate error rates in DF (Marsico, 2004) -lack of reporting of errors/ bias (Jones and Vidalis, 2019) - inability to distinguish tool errors from user errors (Hughes and Karabiyik, 2020; Horsman, 2019a)
Expert skills	-requirements for DF specialists vary amongst jurisdictions (Henseler and van Loenhout, 2018; Kwakman et al., 2011) -competence and impartiality is often assumed by the court (Gross and Mnookin, 2003; Edmond, 2016)

to reject such expert testimony because it was not "derived by the scientific method."[32]

This view is a misinterpretation of Daubert's objectives and does not account for the judge's pivotal role in scrutinizing the use of science for court proceedings. The most important statement of the Dauber court was that in a "case involving scientific evidence, evidentiary reliability will be based upon scientific validity." In fact, the court clarified further in the *Kumho* case that the developed criteria must not be used as a checklist, but as guidance for improvement and an emphasis the importance of scrutinizing scientific evidence by courts. However, Daubert indeed can be interpreted as giving the trial court too broad a discretion to access expert evidence reliability. Moreover, if the field of DF is currently suffering a lack of standards practices and methodologies, how can judges evaluate them in court?

In new and fast developing disciplines like digital forensics, the court is indeed ill-equipped to perform complex reliability tests when the forensic opinion is based on ongoing research. However, research studies show that technology-assisted evidence production is embraced by both judges (Edmond and Roberts, 2011) and prosecuting authorities alike [(Doyle, 2019), Ch. 7]. The enhanced reliance on digital data and automated tools is not accompanied by procedures to scrutinize innovation or by suspect and defendant rights to deal with machine and human bias. Jasanoff emphasizes that "judges cannot surrender to scientists their responsibilities as gatekeepers of evidence, nor can they insist on impossibly high standards of scientific rigor".

The conclusion can be drawn that judges have their important role in verifying the forensic evidence reliability, but they cannot and must not perform scientific validation of digital forensic methods and tools. Their important role is to define the boundaries of permissible expert testimony in court (Gross and Mnookin, 2003) for a particular case. Such validation must be performed prior to the court proceedings. Most importantly, it needs to be done in a standardized and formalized process sufficiently documented to be verifiable in later court proceedings. Standards and formal procedures help forensic scientists to report their scientific findings clearly, allowing judges to understand and evaluate them in a routine manner. More importantly, standards prevent unequal treatment of suspects and defendants, by enforcing objectivity and reducing subjectivity in both examiners and judges.

Often in common law systems, a reliability standard is proposed as an exclusionary rule for low-quality, computer-generated information (Teppler, 2014), while in civil law systems the evidence will be admitted but the question of its reliability will be addressed to its probative weight. The Law Commission in the United Kingdom proposed a "new reliability regime … designed to enable lawyers and judges 'to properly investigate and determine reliability' as part of an improved admissibility practice" in pre-trial proceedings (Edmond, 2012). To the contrary, Edmond argues that "reliability-based admissibility standard for expert opinion evidence, even in conjunction with provision for recourse to court-appointed experts, is unlikely to generate the kinds of changes required to improve the quality of incriminating forensic science". The author considers that often the court is not presented with sufficient information on the data bases used, methodology and testing in order to evaluate the reliability of every piece of evidence in depth [(Edmond, 2012), pp. 55–56]. Judges may prefer to look at previous decisions, rather than applying complex reliability tests. Moreover, forensic experts might stop improving certain procedures and focus on others which are known as accepted by the court. Further proposed solution is evidence reliability to be evaluated by experts before it reaches court. Moreover, if digital forensics

[32] http://www.toxictorts.com/index.php/about-us/articles/62-the-demise-of-daubert-in-state-courts.html.

must be established as a forensic science branch it must not rely on trials to determine the validity of its methods and tools. Therefore, legal, and scientific evaluation of validity must be cumulative, and not subtractive for evidence in court proceedings. The forensic examiner is forming hypotheses about digital facts and tries to falsify them.[33] To the contrary, when evaluating forensic evidence the judge is acting as a "*verificationist*", who is using the instrument of doubt to detect logical inconsistencies with respect to all the evidence in the case and the prosecution hypothesis [(Ross and Thaman, 2018), p. 84 ref. Iacoviello]. The digital artefacts are checked on trial through the legal reasoning about them, which does not provide clear standards of factual accuracy or reliability.

Therefore, the Daubert-similar criteria must be implemented in practice as a formal validation procedure for digital forensics. Data for validation and reliability assessments must be generated as a part of the digital forensic process, and must be done in a practical and expedited way, in order that the court be able to further examine the relevance and reliability of the expert evidence in the legal fact-finding process.

5.2. *Reliability validation during investigation*

There have been several propositions from academia and standardization bodies for a reliability evaluation of digital forensics methods and digital evidence during investigations. However, none of the propositions is related to the implementation of a reliability standard and generating reliability documentation in the everyday work of practitioners.

5.2.1. *Expert commissions*

Edmond and Roberts examine the dangers to the PI in the absence of "effective procedures that lead to the exclusion of incriminating techniques and opinions that are not *demonstrably reliable*" (Edmond and Roberts, 2011; Edmond, 2012). This is the argument that taking the relevance and probative value of expert opinion as default, does not meet the requirement for procedural accuracy which PI has to enforce in order for the standard and burden of proof to be satisfied. It is their central argument that the standard and burden of proof can and do fail to protect innocent defendants, if not supported by procedures for error mitigation throughout the whole investigation. Therefore, they argue that "where the state has a range of procedural options, and the resources are such that any of those could be adopted, obligation to choose the one which will produce the most accurate results."

Based on these considerations, the authors propose a multidisciplinary advisory board to give opinions on the reliability of expert evidence, and to assist judges in their further evaluation. A similar forensic science oversight commission was proposed by Findley (2011), not only to filter out unreliable scientific evidence, but also to ensure that it would be challenged on valid and impartial grounds by the defence. Doyle (2019) examined the lack of impartiality in forensic labs which are either part of or financially dependent upon law enforcement. He proposed the accreditation of a government-funded organization to perform reviews of forensic results for the defence. Further, Broeders et al. (2017) recently proposed that big data analytics and algorithms used in law enforcement "must be made subject to external review by the relevant oversight authority," to scrutinize the quality and use of the data and methods employed. He further argued for more accountability and judicial review of analytical computational methods. However, the author does not develop criteria for how judges can effectively scrutinize such technology. Moreover, some of the analytical tools require reviewers with a multidisciplinary background e.g. computer science and artificial intelligence, or digital forensics and data science.

In Europe, there were also propositions for a European register of judicial experts (Kwakman et al., 2011). It is questionable, however, whether such a specialist body would be a practical authority on Raz's view, [34] ensuring adequate error mitigation and protection for innocent defendants.

An advisory panel will increase bureaucracy and formal compliance paperwork. Its purpose is further described as a "review of available knowledge base" but it is questionable whether current digital forensic techniques would produce the required reliability assessment information. The NRC (2009: 189) report, stated that little 'systematic research has been done to validate the basic premises and techniques" in forensics. The panel should only give an opinion to the judge at trial, but this could strengthen the expert opinion to the extend where the judge's evaluation becomes a formality. Since this solution is resource demanding, it could be used in more complex digital forensic cases. The remaining unsolved issues are how to ensure an effective reliability assessment in every criminal case, what the time limits would be, and how this would deal with the increase of data in every case.

The strange tendency to appoint specialist oversight bodies for criminal procedures which have suffered from deficits for years, also contradicts efforts in many countries to achieve procedural economy and efficiency given the resource constraints in the investigation. For example, *Boyne* reports that given the heavy caseloads in both the US and Germany, investigations are shortened, more confession agreements or plea bargains are utilized, while minor cases are often dismissed (Boyne, 2016). It is expected that in the digital domain the amount of data and number of devices for investigations will drastically increase in the coming years, while the deficit for IT security and digital forensics specialists will deepen. Commission revisions in more complex cases might be a good solution, but this is redundant, costly, burdensome for practitioners and bureaucratic on a large scale.

Furthermore, we should not underestimate the utility of technology to produce accountable and more transparent processes for proving reliability without oversight bodies. The work for the experts will be shifted towards improvement of

[33] Karl Popper, Lecture notes, 1953. Science: Conjectures and refutations, http://www.nemenmanlab.org/~ilya/images/0/07/Popper-1953.pdf, 1953. Visited 2021-05-05.

[34] Joseph Raz, *Between Authority and Interpretation: On the Theory of Law and Practical Reason* (Oxford University Press, 2009. Raz's idea of a theoretical authority might be useful here. The word of a theoretical authority provides us with a reason to hold some belief. The directives of a practical authority, on the other hand, provide us with reasons to perform some act.

reliability and quality assurance processes. A report examining national solutions for forensic science quality assurance shows extensive reliance on formal requirements for qualifications of experts, while the auditing of the quality of the forensic reports and the production of more reliability information during evidence examination is left to the competence of the experts (Kwakman et al., 2011).

5.2.2. *Judicial oversight in investigations*

If we consider that current investigations are more complex given the use of technology, *Risinger* makes two appealing propositions, which can be outlined as appropriate for the digital domain (Risinger and Risinger, 2021). First, he suggests judicial oversight of investigations. This could be important in respect to the cross-border nature of digital evidence, and the increased number of cases where accuracy evaluations must include some assessment of the intrusiveness of investigation measures. A complex test for proportionality, subsidiarity, necessity or different investigation strategies are not guaranteed by a single warrant but require assurance over the whole investigation. Moreover, such judicial oversight could be a separate career development, which includes specific multidisciplinary education, especially in terms of tech-assisted evidence processes. However, the argument could be seen as contrary to the separation between judicial and investigative powers. A middle ground could be similar to the French procedure. The French courts have the possibility to appoint several experts in more complex cases and give the parties a series of opportunities to challenge the opinion of the expert (Delmas-Marty and Spencer, 2005). In the digital investigation process, this could be a combination of experts in data protection, IT security and digital forensics, who observe the accuracy and error mitigation of specific methodologies and give guidelines for standard procedures.

5.2.3. *Crime pattern detection specialists*

The second important point by *Risinger* is the separation of undercover operations that develop information concerning the existence of a crime from the investigation of perpetration in ordinary crimes. A lot of IT specialists and penetration testers are joining efforts to develop crime pattern detection mechanisms and sharing platforms. This could be an adequate solution with respect to computation-assisted investigations and the blurring of lines between forensics and criminology, or security intelligence vs. investigative measures, which are typical of the digital domain. As separate career paths, the procedures for detecting and extrapolating data about criminal activities from big data sets could be evaluated and verified differently, relying on intelligence sharing and extensive cooperation, while the usual investigations could be more rigorous towards individualized evidence procedures, source verification, and relating the data to a concrete suspect or crime. Both propositions require further research, especially with respect to extensive undercover work in dark web spaces and the lack of any standards of procedures there.

5.2.4. *Expert hot tubbing*

Propositions for "devil's advocate" case review and "experts hot-tubbing" (Rares, 2011; Ross, 2013; Sommer, 2009) on pretrial can be also questioned as to their effectiveness, since they require more experts, funding, and time. Expert impartiality in such cases may be questionable due to the small guild numbers in digital forensics and the quality of the reports may vary hugely. Further, defence strategies may be disclosed prematurely in pre-trial expert questioning. Some questions about digital artefacts require legal and multidisciplinary evaluation, which may induce judges to take a passive role in the cross-examination process.

In some countries, it has also been accepted that an expert can demonstrate competence and scientific soundness of her/his method by referring to peer-reviewed articles (Edmond et al., 2010) or previous work. However, the court will find it difficult to make a scientific validation of the quality of such articles and journals. Peer-reviewed methodologies require further validation of the correct application of such methodology to the case at hand, and documentation of divergence or errors.

5.2.5. *Accreditation of defence experts*

Evidence shows that judges currently prefer to base admissibility decisions upon traditional indicia such as formal qualifications and experience. (Edmond and Roberts, 2011) *Doyle* emphasises that currently all forensic experts are part of, or financially dependent on law enforcement agencies (Doyle, 2019). This raises questions about equality of arms and impartiality amongst experts. He proposed accreditation of defence forensic examiners, which "needs to be recognized in the provision of legal aid so that small organizations can afford to acquire and maintain accreditation." However, such a solution is appropriate only for common law countries and will not solve the lack of suspect and defence rights to cross-examine and challenge the employment of technology in investigations. Moreover, in the digital domain the resource constraints are high and the volumes and complexity of data and technology are increasing. Having two digital forensic examinations (for the prosecution and for the defence) goes against procedural efficiency and is impossible in terms of resources, time, or trained personnel.

Therefore, there is a need for further research into the active participation of the defence during the digital forensic examination during the investigation, e.g. access to the chain of custody, the right to use the same DF tools/methods to collect exculpatory evidence, the right to ask the DF examiner questions, and to request scientific validation of the findings.

Reliability validation during investigations is the most viable and appropriate solution for the digital evidence domain. However, digital forensics investigators lack procedures in place to produce the information necessary for reliability validation. Legislation does not enforce such reliability assessment and standard validation procedures, which results in routine admission of digital evidence that does not meet forensic evidence standards. The criteria identified in *Section D.II* Table 1 could serve for further development of validation procedures in order to overcome these limitations in practice.

6. Conclusions and further work

This paper has examined the emerging challenges to the right to a fair trial at the investigation stage due to *(i)* the inappro-

priate and inconsistent use of technology; *(ii)* old procedural guarantees, which are not adapted to contemporary digital evidence processes and services; *(iii)* the lack of reliability testing in digital forensics practice. It was argued that PI has a role to strengthen the evidence at the investigation stage, where data (information) is heavily processed and corroborated in growing amounts. Any use of technology during the investigation must meet a reliability standard which should be weighed against the probability of wrongful conviction and error.

The DF challenges taxonomy proves that the way digital forensic "science" is performed in practice is led by operational and investigative objectives, which lacks scientific validity and process quality assurance. Digital forensics has no established procedures to produce the information necessary for reliability validation. Legislation is also not enforcing such reliability assessments which results in routine admission of digital evidence without meeting digital forensics standards. Therefore, it is further considered that the employment of computational methods in investigations requires a level of accountability and transparency which can facilitate an evaluation of the proportionality and reliability of the method and better ensure the position of suspects/defendants with respect to the technology used and results in digital forensics. The amount of data and number of devices for investigations will drastically increase over the coming years, while the deficit for digital forensics specialists will deepen. Commission revisions of DF methodology in more complex cases might be a good solution, but it is redundant, costly, burdensome for practitioners and bureaucratic on a large scale. Automated and standardized reliability evaluation on pre-trial during the DF procedure, combined with formal verification of the DF procedure by the court must be considered as the most suitable method for tackling the current reliability validation crisis in DF (Tables 1 and 2).

The presumption of innocence must be taken into consideration in further development of the reliability requirements for investigative technology, specifically focused on the further processing of such data by automated means, the protection of newly-inferred data, and the issue of repurposing and merging of data from different data bases and sources with different levels of accuracy. Moreover, the reliability and fairness of the use of technology must be evaluated atomistic – at each stage of the processing. More importantly, holistic evaluation of periodical or analytical combinations of investigative techniques must meet the same standard. The criteria developed in *Section D.II* Table 1 can serve for further development of formal validation procedures in digital forensics.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

1. No data was used for the research described in the article.

Acknowledgments

This work was supported by the European Union's Horizon 2020 Research and Innovation Programme under the Marie Skłodowska-Curie Grant Agreement No. 722482.

REFERENCES

Adams R, Hobbs V, Mann G. The Advanced Data Acquisition Model (Adam): a process model for digital forensic practice. J. Digit. Forensics Secur. Law 2013;8(4) Jan. doi:10.15394/jdfsl.2013.1154.

Alendal G, Dyrkolbotn GO, Axelsson S. Forensics acquisition — analysis and circumvention of Samsung secure boot enforced common criteria mode. Digit. Investig. 2018;24:S60–7 Mar. doi:10.1016/j.diin.2018.01.008.

Amuchi F, Al-Nemrat A, Alazab M, Layton R. Identifying cyber predators through forensic authorship analysis of chat logs. Proceedings of the Third Cybercrime and Trustworthy Computing Workshop; 2012. p. 28–37 Oct.

Årnes A. Digital Forensics: An Academic Introduction (Ed.). Hoboken, NJ: John Wiley & Sons Inc; 2018.

Arshad H, Jantan AB, Abiodun OI. Digital forensics: review of issues in scientific validation of digital evidence. J. Inf. Process. Syst. 2018;14(2):346–76 Apr.

Ashworth A. Four Threats to the Presumption of Innocence. Int. J. Evid. Proof 2006;10(4):241–79 Jul. doi:10.1350/ijep.10.4.241.

Beebe N. Digital forensic research: the good, the bad and the unaddressed. Proceedings of the Advances in Digital Forensics V; 2009. p. 17–36 Jan.

Boister N. An Introduction to Transnational Criminal Law. Second edition. Oxford, United Kingdom: Oxford University Press; 2018.

Boyne SM. Procedural economy in pre-trial procedure: developments in Germany and the United States. Comp. Crim. Proced. 2016. Jun.Accessed: Feb. 18, 2020. [Online]. Available https://www.elgaronline.com/view/edcoll/9781781007181/9781781007181.00016.xml.

Broeders D, Schrijvers E, van der Sloot B, van Brakel R, de Hoog J, Hirsch Ballin E. Big data and security policies: towards a framework for regulating the phases of analytics and use of big data. Comput. Law Secur. Rev. 2017;33(3):309–23 Jun. doi:10.1016/j.clsr.2017.03.002.

Brown I, Korff D. Terrorism and the proportionality of internet surveillance. Eur. J. Criminol. 2009;6(2):119–34 Mar. doi:10.1177/1477370808100541.

Budish R, Burkert H, Gasser U. Encryption Policy and Its International Impacts: A Framework for Understanding Extraterritorial Ripple Effects. A Hoover Institution Essay; 2018 Aegis Series Paper No. 1804Accessed: Sep. 19, 2020. [Online]. Available https://dash.harvard.edu/handle/1/36291726.

Callen, C. (2015)."Human deliberation in fact-finding and human rights in the law of evidence |," https://lawexplores.com/human-deliberation-in-fact-finding-and-human-rights-in-the-law-of-evidence/ (accessed Jul. 20, 2020).

Carrier B. Open Source Digital Forensics Tools: The Legal Argument; 2002.

Carrier B. An event-based digital forensic investigation framework. Presented at the Digital Forensic Research Workshop, Baltimore, 2004.

Carrier BD. A hypothesis-based approach to digital forensic investigations. Theses Diss. Available ProQuest 2006:1–169 Jan.

Casey E. What does 'forensically sound' really mean? Digit. Investig. 2007;4:49–50 Jun. doi:10/bh3dcs.

Casey E. What does 'forensically sound' really mean? Digit. Investig. 2007;4(2):49–50 Jun. doi:10.1016/j.diin.2007.05.001.
Casey E. The chequered past and risky future of digital forensics. Aust. J. Forensic Sci. 2019;51(6):649–64 Nov. doi:10.1080/00450618.2018.1554090.
Ciraco D. Reverse engineering. Windsor Rev. Leg. Soc. Issues 2001;11:41–72.
Council of the European Union. Draft Council Conclusions on the Way Forward in View of the Creation of An European Forensic Science Area. Council of the European Union; 2016 Feb. 18Accessed: Mar. 27, 2018. [Online]. Available http://data.consilium.europa.eu/doc/document/ST-6078-2016-INIT/en/pdf.
De Hert PJA. Balancing security and liberty within the European human rights framework. A critical reading of the Court's case law in the light of surveillance and criminal law enforcement strategies after 9/11. Utrecht Law Rev. 2005;1(1):68 Jul. doi:10.18352/ulr.4.
Delmas-Marty M, Spencer JR. European Criminal Procedures. 1st pbk ed Eds. Cambridge ; New York: Cambridge University Press; 2005.
Doyle S. Quality Management in Forensic Science. London, United Kingdom ; San Diego, CA, United States: Elsevier : Academic Press, is an imprint of Elsevier; 2019.
Dror IE. Cognitive and human factors in expert decision making: six fallacies and the eight sources of bias. Anal. Chem. 2020;92(12):7998–8004 Jun. doi:10.1021/acs.analchem.0c00704.
Edmond G, Roberts A. Procedural fairness, the criminal trial and forensic science and medicine. Forensic Sci. Med. 2011;33(3):36.
Edmond G, et al. Atkins v the emperor: the 'cautious' use of unreliable 'expert' opinion. Int. J. Evid. Proof 2010;14(2):146–66 Apr. doi:10.1350/ijep.2010.14.2.349.
Edmond G. The admissibility of incriminating expert opinion evidence in the US, England and Canada. Judic. Off. Bull. 2011;23:67–70 Sep.
Edmond G. Is reliability sufficient? The law commission and expert evidence in international and interdisciplinary perspective (Part 1). Int. J. Evid. Proof 2012;16:30–65 Jan. doi:10.1350/ijep.2012.16.1.391.
Edmond G. Advice for the courts? Sufficiently reliable assistance with forensic science and medicine (Part 2). Int. J. Evid. Proof 2012;16(3):263–97 Jul. doi:10.1350/ijep.2012.16.3.405.
Edmond G. Legal versus non-legal approaches to forensic science evidence. Int. J. Evid. Proof 2016;20(1):3–28 Jan. doi:10.1177/1365712715613470.
European Commission. E-evidence - Cross-Border Access to Electronic Evidence. European Commission; 2017 European Commission https://ec.europa.eu/info/strategy/justice-and-fundamental-rights/criminal-justice/e-evidence_en (accessed Apr. 18, 2018).
Findley KA, Scott MS. The Multiple Dimensions of Tunnel Vision in Criminal Cases. Rochester, NY: Social Science Research Network; 2006 SSRN Scholarly Paper ID 911240Jun. Accessed: Aug. 11, 2020. [Online]. Available https://papers.ssrn.com/abstract=911240.
Findley K. Innocents at risk: adversary imbalance, forensic science, and the search for truth. Seton Hall Law Rev. 2011;38(3). Nov.[Online]. Available https://scholarship.shu.edu/shlr/vol38/iss3/7.
Foster KR, Huber PW. Judging Science: Scientific Knowledge and the Federal Courts. Cambridge, Mass: MIT Press; 1999 1. paperback ed.
Friheim I. Practical Use of Dual Tool Verification in Computer Forensics. University College Dublin; 2016.
Garfinkel S, Farrell P, Roussev V, Dinolt G. Bringing science to digital forensics with standardized forensic corpora. Digit. Investig. 2009;6:S2–S11 Sep. doi:10.1016/j.diin.2009.06.016.
Gerber M, Leeson J. Formalization of computer input and output: the Hadley model. Digit. Investig. 2004;1(3):214–24 Sep. doi:10.1016/j.diin.2004.07.001.
Gless S. Truth or due process? The use of illegally gathered evidence in criminal trials - Germany. SSRN Electron. J. 2010. doi:10.2139/ssrn.1743530.
Gless S. Transnational cooperation in criminal matters and the guarantee of a fair trial: approaches to a general principle. Utrecht Law Rev. 2013;9(4):90 Sep. doi:10.18352/ulr.244.
Gross, S., & Mnookin, J. (2003)."Expert information and expert evidence: a preliminary taxonomy," Articles, Jan., [Online]. Available: https://repository.law.umich.edu/articles/570
Gross, S. (1996)."The risks of death: why erroneous convictions are common in capital cases (Symposium: The New York Death Penalty in Context)," Articles, Jan., [Online]. Available: https://repository.law.umich.edu/articles/193
Hadjimatheou K. Surveillance technologies, wrongful criminalisation, and the presumption of innocence. Philos. Technol. 2017;30(1):39–54 Mar. doi:10.1007/s13347-016-0218-2.
Hay R. fastboot oem vuln: android bootloader vulnerabilities in vendor customizations. Presented at the 11th {USENIX} Workshop on Offensive Technologies ({WOOT} 17), 2017. https://www.usenix.org/conference/woot17/workshop-program/presentation/hay.
Henseler H, van Loenhout S. Educating judges, prosecutors and lawyers in the use of digital forensic experts. Digit. Investig. 2018;24:S76–82 Mar. doi:10.1016/j.diin.2018.01.010.
Hildebrandt M. Criminal law and technology in a data-driven society. In: Dubber MD, Hörnle T, editors. The Oxford Handbook of Criminal Law. Oxford University Press; 2014.
Hong I, Yu H, Lee S, Lee K. A new triage model conforming to the needs of selective search and seizure of electronic evidence. Digit. Investig. 2013;10(2):175–92 Sep. doi:10.1016/j.diin.2013.01.003.
Horsman G. Framework for reliable experimental design (FRED): a research framework to ensure the dependable interpretation of digital data for digital forensics. Comput. Secur. 2018a;73:294–306 Mar. doi:10/gcx6dd.
Horsman G. Framework for Reliable Experimental Design (FRED): A research framework to ensure the dependable interpretation of digital data for digital forensics. Comput. Secur. 2018b;73:294–306 Mar. doi:10.1016/j.cose.2017.11.009.
Horsman G. Tool testing and reliability issues in the field of digital forensics. Digit. Investig. 2019a;28:163–75 Mar. doi:10.1016/j.diin.2019.01.009.
Horsman G. Formalising investigative decision making in digital forensics: Proposing the Digital Evidence Reporting and Decision Support (DERDS) framework. Digit. Investig. 2019b;28:146–51 Mar. doi:10.1016/j.diin.2019.01.007.
Hughes N, Karabiyik U. Towards reliable digital forensics investigations through measurement science. WIREs Forensic Sci. 2020:e1367 vol. n/a, no. n/a. doi:10.1002/wfs2.1367.
ISO, I. "ISO/IEC 27037 eForensics guidelines for identification, collection, acquisition and preservation of digital evidence," (2012). https://www.iso27001security.com/html/27037.html (accessed Sep. 03, 2020).
Ieong RSC. FORZA – Digital forensics investigation framework that incorporate legal issues. Digit. Investig. 2006;3:29–36 Sep. doi:10/cm84cg.
Igel C. No free lunch theorems: limitations and perspectives of metaheuristics. In: Borenstein Y, Moraglio A, editors. Theory and Principled Methods for the Design of Metaheuristics. Berlin, Heidelberg: Springer; 2014. p. 1–23.
Innocence Project, *Innocence project*. https://www.innocenceproject.org/ (accessed Jan. 29, (2020)).
Jackson JD, Summers SJ. The Internationalisation of Criminal Evidence: Beyond the Common Law and Civil Law Traditions. Cambridge University Press; 2012.

Jasanoff S. Law's knowledge: Science for justice in legal settings. Am. J. Public Health 2005;95(Suppl 1):S49–58. doi:10.2105/AJPH.2004.045732.
Jones A, Vidalis S. Rethinking digital forensics. Ann. Emerg. Technol. Comput. 2019;3:41–53 Apr. doi:10.33166/AETiC.2019.02.005.
Kloosterman A, et al. The interface between forensic science and technology: How technology could cause a paradigm shift in the role of forensic institutes in the criminal justice system. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 2015;370(1674) Aug. doi:10.1098/rstb.2014.0264.
Kohn MD, Eloff MM, Eloff JHP. Integrated digital forensic process model. Comput. Secur. 2013;38:103–15 Oct. doi:10/f5gxzk.
Koops B-J, Kosta E. Looking for some light through the lens of 'cryptowar' history: Policy options for law enforcement authorities against 'going dark. Comput. Law Secur. Rev. 2018;34(4):890–900 Aug. doi:10.1016/j.clsr.2018.06.003.
Kusak M. Common EU minimum standards for enhancing mutual admissibility of evidence gathered in criminal matters. Eur. J. Crim. Policy Res. 2017;23(3):337–52 Sep. doi:10.1007/s10610-017-9339-0.
Kwakman, N. J. M., Nijboer, J. A., Keulen, B. F., & Elzinga, H. K. (2011)."Expert registers in criminal cases. Governance in criminal proceedings.", Accessed: Jun. 25, 2020. [Online]. Available: https://www.rug.nl/rechten/congressen/archief/2011/governancemeetslaw/workingpapers/papernijboerkeulen.pdf
Layton R, Watters P, Dazeley R. Authorship attribution for twitter in 140 characters or less. Proceedings of the Second Cybercrime and Trustworthy Computing Workshop; 2010. p. 1–8 Jul.
Lillis D, Breitinger F, Scanlon M. Hierarchical bloom filter trees for approximate matching. J. Digit. Forensics Secur. Law 2018;13(1) Mar. doi:10.15394/jdfsl.2018.1489.
Lyle JR. If error rate is such a simple concept, why don't I have one for my forensic tool yet? Digit. Investig. 2010;7:S135–9 Aug. doi:10.1016/j.diin.2010.05.017.
Marshall, A., & Paige, R. (2018).*Requirements in digital forensics method definition: observations from a UK study*.
Marsico CV. CERIAS Tech Report: Computer Evidence v. Daubert: The Coming Conflict. Purdue University School of Technology; 2004 [Online]. Available https://www.cerias.purdue.edu/assets/pdf/bibtex_archive/2005-17.pdf.
Mason S, Seng D. Electronic Evidence. Fourth. University of London, Institute of Advanced Legal Studies; 2017.
Mifsud Bonnici JP, Tudorica M, Cannataci JA. The European legal framework on electronic evidence: complex and in need of reform. In: Biasiotti MA, Cannataci J, Mifsud Bonnici JP, Turchi F, editors. Handling and Exchanging Electronic Evidence Across Europe. Cham: Springer International Publishing : Imprint: Springer; 2018 2018.
Milaj J, Mifsud Bonnici JP. Unwitting subjects of surveillance and the presumption of innocence. Comput. Law Secur. Rev. 2014;30(4):419–28 Aug. doi:10.1016/j.clsr.2014.05.009.
Mocas S. Building theoretical underpinnings for digital forensics research. Digit. Investig. 2004;1(1):61–8 Feb. doi:10.1016/j.diin.2003.12.004.
Montasari R. A comprehensive digital forensic investigation process model. Int. J. Electron. Secur. Digit. Forensics 2016;8(4):285–302 Jan. doi:10/gcx6ds.
Morgan RM. Conceptualising forensic science and forensic reconstruction. Part II: the critical interaction between research, policy/law and practice. Sci. Justice 2017;57(6):460–7 Nov.doi: 10/gcqq93.
Nguyen HT, Franke K, Petrovic S. Towards a generic feature-selection measure for intrusion detection. Proceedings of the 20th International Conference on Pattern Recognition; 2010. p. 1529–32.
Page H, Horsman G, Sarna A, Foster J. A review of quality procedures in the UK forensic sciences: What can the field of digital forensics learn? Sci. Justice 2019;59(1):83–92 Jan. doi:10.1016/j.scijus.2018.09.006.
Patel A, ó Ciardhuaéin S. Impact of forensic computing on telecommunications. IEEE Commun. Mag. 2000;38(11):64–7. doi:10.1109/35.883490.
PCAST Releases report on forensic science in criminal courts | whitehouse.gov." https://obamawhitehouse.archives.gov/blog/2016/09/20/pcast-releases-report-forensic-science-criminal-courts (accessed Mar. 06, (2020)).
Pollitt M, Casey E, Jaquet-Chiffelle DO, Gladyshev P. A Framework for Harmonizing Forensic Science Practices and Digital/Multimedia Evidence. NIST - OSAC Task Group on Digital/Multimedia Science; 2018 Jan. 11 https://www.nist.gov/news-events/news/2018/01/framework-harmonizing-forensic-science-practices-and-digitalmultimedia (accessed Jul. 02, 2020).
Rares S. Using the 'hot tub': how concurrent expert evidence aids understanding of issues. Judic. Rev. 2011:171–86 Apr.
Reedy P. Interpol review of digital evidence 2016 - 2019. Forensic Sci. Int. Synergy 2020 Mar. doi:10.1016/j.fsisyn.2020.01.015.
Risinger, D. M., & Risinger, L. C. (2021)."Innocence is different: taking innocence into account in reforming criminal procedure," vol. 56, p. 41
Risinger DM, Saks MJ, Thompson WC, Rosenthal R. The Daubert/Kumho implications of observer effects in forensic science: hidden problems of expectation and suggestion. Calif. Law Rev. 2002. doi:10.2307/3481305.
Risinger DM. Navigating Expert Reliability: Are Criminal Standards of Certainty Being Left on the Dock?. Rochester, NY: Social Science Research Network; 2000 SSRN Scholarly Paper ID 251033Accessed: Jun. 25, 2020. [Online]. Available https://papers.ssrn.com/abstract=251033.
Risinger DM. Guilt vs. Guiltiness: Are the Right Rules for Trying Factual Innocence Inevitably the Wrong Rules for Trying Culpability?. Rochester, NY: Social Science Research Network; 2008 SSRN Scholarly Paper ID 1359948Accessed: Aug. 11, 2020. [Online]. Available https://papers.ssrn.com/abstract=1359948.
Risinger D. The five functions of forensic science and the validation issues they raise: a piece to incite discussion on validation. Seton Hall Law Rev. 2018;48(3). Jun.[Online]. Available https://scholarship.shu.edu/shlr/vol48/iss3/6.
Ross JE, Thaman S. Comparative Criminal Procedure. Paperback edition Eds. Cheltenham, UK Northampton, MA, USA: Edward Elgar Publishing; 2018.
Ross, A. (2013)."Murky waters: an expert's perspective on the effectiveness of expert conclaves and 'hot tubs,'" *Preced. Syd. NSW No 119*, pp. 30–34,
Rumold M. Playpen: the Story of the FBI's Unprecedented and Illegal Hacking Operation. Electronic Frontier Foundation; 2016 https://www.eff.org/deeplinks/2016/09/playpen-story-fbis-unprecedented-and-illegal-hacking-operation%20 (accessed Aug. 12, 2020).
Saks M, Faigman DL. Failed forensics: how forensic science lost its way and how it might yet find it. Annu. Rev. Law Soc. Sci. 2008;4:149–71. doi:10.1146/annurev.lawsocsci.4.110707.172303.
Saks MJ, Koehler JJ. The coming paradigm shift in forensic identification science. Science 2005;309(5736):892–5 Aug. doi:10.1126/science.1111565.
Schum DA. The Evidential Foundations of Probabilistic Reasoning. Evanston, Ill: Northwestern University Press; 2001.
Shapiro BJ. A Culture of Fact: England, 1550–1720. Ithaca, NY: Cornell Univ. Press; 2003 1. paperback printing.
SKAPP. Daubert: The Most Influential Supreme Court Ruling You've Never Heard Of; 2003.

Sliedregt EV. A contemporary reflection on the presumption of innocence. Rev. Int. Droit Penal 2009;80(1):247–67.
Sommer P. Meetings between experts: a route to simpler, fairer trials? Digit. Investig. 2009;5(3–4):146–52 Mar. doi:10.1016/j.diin.2008.11.002.
Sommer P. Forensic science standards in fast-changing environments. Sci. Justice J. Forensic Sci. Soc. 2010;50:12–17 Mar. doi:10.1016/j.scijus.2009.11.006.
Sremack JC. The gap between theory and practice in digital forensics. Presented at the Conference on Digital Forensics, Security and Law, 2007. https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.823.8791&rep=rep1&type=pdf.
Stein A. Foundations of Evidence Law. Oxford ; New York: Oxford University Press; 2005.
Stoykova R, Franke K. Standard representation for digital forensic processing. Proceedings of the 13th International Conference on Systematic Approaches to Digital Forensic Engineering (SADFE); 2020. p. 46–56 May.
Street C. Streaming the international silver platter doctrine: coordinating transnational law enforcement in the age of global terrorism and technology. Columbia J. Transnatl. Law 2011;45(2). [Online]. Available http://blogs2.law.columbia.edu/jtl/streaming-the-international-silver-platter-doctrine-coordinating-transnational-law-enforcement-in-the-age-of-global-terrorism-and-technology/.
Stuckenberg C-F. *Untersuchungen zur Unschuldsvermutung*: Berlin. Boston: De Gruyter; 1998.
Sunde N, Dror IE. Cognitive and human factors in digital forensics: problems, challenges, and the way forward. Digit. Investig. 2019;29:101–8 Jun. doi:10.1016/j.diin.2019.03.011.
T. F. The UK National Police Chiefs Council, "Digital forensic science strategy." Jul. (2020)., [Online]. Available: https://www.npcc.police.uk/Digital%20Forensic%20Science%20Strategy%202020.pdf
Teppler Steven. Testable reliability: a modernized approach to ESI admissibility. Ave Maria Law Rev. 2014;12(2):213.
Trechsel S, Summers SJ. Human Rights in Criminal Proceedings. Oxford: Oxford Univ. Press; 2006.
Tully G, Cohen N, Compton D, Davies G, Isbell R, Watson T. Quality standards for digital forensics: learning from experience in England & Wales. Forensic Sci. Int. Digit. Investig. 2020;32 Mar. doi:10.1016/j.fsidi.2020.200905.
Twining W. Rethinking Evidence: Exploratory Essays. Cambridge; New York: Cambridge University Press; 2006.
Ulges A, Stahl A. Automatic detection of child pornography using color visual words. Proceedings of the IEEE International Conference on Multimedia and Expo; 2011. p. 1–6 Jul.
United Nations Office on Drugs and Crime (UNODC), "Draft comprehensive study on cybercrime," (2013). https://www.unodc.org/e4j/data/_university_uni_/draft_comprehensive_study_on_cybercrime.html?lng=en (accessed May 05, 2021).
van Baar RB, van Beek HMA, van Eijk EJ. Digital forensics as a service: a game changer. Digit. Investig. 2014;11:S54–62 May. doi:10.1016/j.diin.2014.03.007.
van Wijk MC. Cross-Border Evidence Gathering: Equality of Arms within the EU?. The Hague, The Netherlands: Eleven International Publishing; 2017.
Vermeulen G, De Bondt W, van Damme Y. EU Cross-Border Gathering and Use of Evidence in Criminal Matters: Towards Mutual Recognition of Investigative Measures and Free Movement of Evidence?. Antwerpen; 2010 Portland: Maklu.
Vervaele JAE. Special procedural measures and the protection of human rights
general report. Utrecht Law Rev. 2009;5(2):66–103 Oct. doi:10.18352/ulr.103.
Vincze EA. Challenges in digital forensics. Police Pract. Res. 2016;17(2):183–94 Mar. doi:10.1080/15614263.2015.1128163.
Vuille J. Admissibility and appraisal of scientific evidence in continental European criminal justice systems: Past, present and future. Aust. J. Forensic Sci. 2013;45(4):389–97 Dec. doi:10.1080/00450618.2012.738248.
Wolpert DH, Macready WG. Coevolutionary free lunches. IEEE Trans. Evol. Comput. 2005;9(6):721–35 Dec. doi:10.1109/TEVC.2005.856205.
Wundram M, Freiling FC, Moch C. Anti-forensics: the next step in digital forensics tool testing. Proceedings of the Seventh International Conference on IT Security Incident Management and IT Forensics; 2013. p. 83–97 Mar.
Zagaris B, Plachta M. Transnational organized crime section I EU and law enforcement dismantle encrypted network of transnational organized crime. Int. Enforc. Law Report. 2020;36(7):248–55.
Zheng R, Li J, Chen H, Huang Z. A framework for authorship identification of online messages: Writing-style features and classification techniques. J. Am. Soc. Inf. Sci. Technol. 2006;57(3):378–93 Feb. doi:10.1002/asi.20316.