

# 2

---

## European Approaches to AI-Generated Evidence in Criminal Proceedings

---

SABINE GLESS\*

### I. Mapping the Field from a European Perspective

In the past, it has been common opinion that whizzes from the United States (US) provide the technology for building complex IT systems, and European specialists the regulation to attempt to govern them.<sup>1</sup> However, the chapters resulting from the Luxembourg CRIM\_AI project presented in this volume highlight that European experts also provide authorities in Europe with novel artificial intelligence (AI) systems and that recent US legislative initiatives in fact aimed to regulate the relevant technology more closely.<sup>2</sup> Furthermore, the country reports individually illustrate how the criminal trials in different jurisdictions and legal families react to evidence generated by (or with the help of) such systems, and whether and how the legal systems are prepared to vet the reliability and trustworthiness of such evidence in a criminal proceeding.<sup>3</sup>

The question of whether existing rules on criminal procedure, in particular evidence law and procedural guarantees, are sufficient to address the specific nature and the associated pitfalls of AI-generated evidence requires a multi-level answer. The various and ambitious European legal frameworks aiming at trustworthy AI are unique and have the potential to assist in the development of a new European approach to regulating the use of AI-generated evidence in criminal proceedings. It warrants a detailed analysis, not only because the different regulations are complex and seem to overlap, but also because they illustrate how insulated and, at times, rather indistinct initiatives can eventually form a strong structure for regulating new technology. A prominent example for this

\* Professor of criminal law and criminal procedural law at the University of Basel. The author wishes to thank Anne Herrmann for her valuable assistance in researching and editing this chapter.

<sup>1</sup> 'The EU Wants to Become the World's Super-Regulator in AI' *The Economist* (24 April 2021), available at [www.economist.com/europe/2021/04/24/the-eu-wants-to-become-the-worlds-super-regulator-in-ai](http://www.economist.com/europe/2021/04/24/the-eu-wants-to-become-the-worlds-super-regulator-in-ai).

<sup>2</sup> Ch 1.IV ff. See also European Committee on Crime Problems (CDPC), *Assessment of the Answers to the Questionnaire on Artificial Intelligence and Criminal Justice* (Council of Europe, 7 November 2019) 11, available at [rm.coe.int/cdpc-2019-17-draft-assessment-of-the-answers-to-the-questionnaire-on-a/168098e24c](http://rm.coe.int/cdpc-2019-17-draft-assessment-of-the-answers-to-the-questionnaire-on-a/168098e24c).

<sup>3</sup> See chs 4–9.

is how the European Convention on Human Rights (ECHR) and other Council of Europe (CoE) instruments, on the one hand, and, on the other, an ambitious EU policy based on the General Data Protection Regulation (GDPR)<sup>4</sup> and the EU AI Act<sup>5</sup> could intertwine when it comes to establishing starting points to conceptualise a framework for a legal analysis of AI-generated evidence.

## A. Stakeholders

Two institutions foster work on using AI systems in criminal proceedings. The work done by the European Union, and in particular the European Commission, as well as the work conducted in the CoE reflect regional approaches that will impact the use of AI systems in criminal justice far beyond the European continent.

However, the most important stakeholders in the European approach to using AI for evidentiary issues are the European states and their authorities, as criminal justice remains a primarily domestic affair. Issues of evidence and proof build on long traditions concerning how to establish truth in a criminal case<sup>6</sup> and the relevant European legislative frameworks do not specifically regulate the evidentiary use of AI. Since 2018 the EU has drawn on the advice of its high-level expert group on AI with the mission of supporting the implementation of the European initiative on AI. This includes elaborating recommendations for future AI-related policy development and on ethical, legal and societal issues related to AI. Not surprisingly, given the division of competences between the European and the domestic level, there is no mandate on AI-generated evidence or even use of AI in criminal justice. Therefore, policy recommendations elaborated by scientific associations that are addressed to national policy-makers provide important insights and will be also included in this chapter, such as the resolutions of the Association International de Droit Pénal.<sup>7</sup>

Nevertheless, in particular with regard to the use of AI systems, European regulation plays an important role. In conformity with its traditional role, the CoE takes a rather bottom-up approach, orchestrated by a close exchange with state parties, whilst the EU follows its own (complex) procedures, involving the relevant EU institutions, for the adoption of supra-national legal instruments with binding force and judicial monitoring.

<sup>4</sup> Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC [2016] OJ L 119/1.

<sup>5</sup> Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 [2024] (EU AI Act).

<sup>6</sup> For an illustration of German and US law, see: S Gless, F Lederer and T Weigend, 'AI-Based Evidence in Criminal Trials?' (2024) 1 *Tulsa Law Review* 1, 23.

<sup>7</sup> cf J Lelieur (ed), *Artificial Intelligence and Administration of Criminal Justice. International Colloquium, Buenos Aires, Argentina, 28–31 March 2023* (Revue Internationale de Droit Pénal, 2023).

### i. Council of Europe

The CoE has been known as a guardian of human rights and democracy since the state parties established the European Court of Human Rights (ECtHR) and several committees safeguarding certain standards regarding human rights, democracy and the rule of law at the end of World War II. The ECHR does not stipulate any specific protection against AI systems as such. But, for instance, in criminal justice fair trial rights such as the right to examine ‘witnesses’ against oneself based on Article 6(3)(d) ECHR or Article 8 ECHR do apply when seeing to the overall standards.<sup>8</sup> It is however important to keep in mind that the case-by-case approach and deference to national legislatures and administrative bodies leave a wide margin of discretion as how to protect human rights in their system.<sup>9</sup>

Different working groups and committees have addressed issues of human rights, democracy and prevention of discrimination in the age of AI. The CoE’s initiatives include the 2018 European Ethical Charter on the use of Artificial Intelligence in judicial systems and their environment.<sup>10</sup> It is the first European text addressing ethical principles relating to such use of AI. It outlines principles to guide policy-makers, legislators and justice professionals in general to not only enhance efficiency and certain quality features in criminal justice systems, but to ensure that the design and implementation of AI tools are compatible with fundamental rights.<sup>11</sup> Here we find the blueprints for the principles of transparency, impartiality and fairness: making data-processing methods accessible and understandable, ensuring ‘user control’, and – also of great importance – authorising external audits. These ideas were developed more broadly later in the EU AI Act, which mixes a product safety approach with judicial rights protection, and the CoE’s Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law. The latter is an umbrella convention for further CoE policy and activity adopted in May 2024 by the Committee of Ministers and open for signature from September 2024 onwards. The role of developing this Convention has been given to a newly established Committee on Artificial Intelligence (CAI).<sup>12</sup>

The CAI aims to ensure an unconstrained application of human rights and the principle of rule of law in situations where AI systems assist or replace human decision-making or perform other tasks relevant in such contexts. Unlike the EU AI Act, however, it carries no explicit provisions on law enforcement use of AI evidence. It is nevertheless

<sup>8</sup> S Gless, ‘AI in the Courtroom: A Comparative Analysis of Machine Evidence in Criminal Trials’ (2020) 51 *Georgetown Journal of International Law* 195, 213.

<sup>9</sup> S Gless and J Martin, ‘The Comparative Method in European Courts: A Comparison Between the CJEU and ECtHR’ (2013) *Bergen Journal of Criminal Law & Criminal Justice* 1, 36, available at [dx.doi.org/10.15845/bjclj.v1i1.522](https://dx.doi.org/10.15845/bjclj.v1i1.522).

<sup>10</sup> European Commission for the Efficiency of Justice, *European Ethical Charter on the use of Artificial Intelligence in judicial systems and their environment* (Council of Europe, 4 December 2018), available at [rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c](https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c).

<sup>11</sup> AD Reiling, ‘Courts and Artificial Intelligence’ (2020) 11 *International Journal of Computer Applications for Court Administration*, available at 1 [doi.org/10.36745/ijca.343](https://doi.org/10.36745/ijca.343).

<sup>12</sup> ‘Committee on Artificial Intelligence (CAI)’ (*Council of Europe*), available at [www.coe.int/en/web/artificial-intelligence/cai](https://www.coe.int/en/web/artificial-intelligence/cai).

understood that it builds on the vast human rights law, including ECtHR case law, which binds Member States to certain standards of transparency, accountability, etc.<sup>13</sup>

Even though the CoE's focus is on safeguarding democratic processes and upholding human rights overall, it acknowledges the fact that AI, developed and used across borders, will also impact criminal justice. The commitment is to develop a common approach to basic principles that should govern how we, as humanity, develop and use AI systems. With this it links to the CoE's commitment on human rights, democracy and the rule of law through standards and its tradition to invite all states – be they a member of the CoE or not – to join the Convention. This opens a promising potential, illustrated by the CoE's Convention on Cybercrime (ETS 185) adopted in Budapest on 23 November 2001 (Budapest Convention) with its protocols. The framework provided many novel options for cross-border investigations on cybercrime but did not, and still does not, specifically address using AI systems. As of now, it seems unclear how the development in technology will affect the state parties' obligation to have an effective international scheme for the collection and preservation of digital evidence.<sup>14</sup>

Yet, if the CAI and the EU AI Act apply in parallel, potential conflicts could arise when EU Member States sign up to CAI, as the two legal frameworks devote themselves to the same principles, such as transparency or human oversight, but might develop different scopes or practices. In particular, where the EU AI Act links product safety with fundamental rights protection unknown consequences for criminal justice may arise in the future.

The Council of Europe's Committee on Criminal Problems (CDPC) – which has dealt with various activities in the field of criminal justice for a long time – once more provided a forum for state parties to identify a new topic for intergovernmental work and elaborate on a need for new conventions when choosing a project on Artificial Intelligence and Criminal Law (see below II.C.i.).

## *ii. European Union*

The EU has been very active in developing a strategy for the digital age. The primary focus on data protection is now coupled<sup>15</sup> with regulations which ensure that AI systems in the EU are safe and respect fundamental rights and values, presenting a proposal for an Artificial Intelligence Liability Directive,<sup>16</sup> adopting the EU AI Act,<sup>17</sup> and constantly communicating its digital strategy.<sup>18</sup>

<sup>13</sup> eg, D Leslie et al, 'Artificial Intelligence, Human Rights, Democracy, and the Rule of Law: A Primer' (Council of Europe, 2021), <https://doi.org/10.48550/arXiv.2202.02776>.

<sup>14</sup> S Tosza, 'All evidence is equal, but electronic evidence is more equal than any other: The relationship between the European Investigation Order and the European Production Order' (2020) 11 *New Journal of European Criminal Law* 161, 168 f.

<sup>15</sup> For more details see ch 3.

<sup>16</sup> European Commission, 'Proposal for a Directive of the European Parliament and of the Council on adapting non-contractual civil liability rules to artificial intelligence' COM(2022) 496 final.

<sup>17</sup> EU AI Act.

<sup>18</sup> 'EU Digital Strategy' (*EU4Digital*), available at [eufordigital.eu/discover-eu/eu-digital-strategy](https://eufordigital.eu/discover-eu/eu-digital-strategy).

It is notable that the European Parliament came up with various initiatives, but eventually the European Commission once more proved the motor of integration.<sup>19</sup>

As has been pointed out, criminal justice remains the domain of Member States on the bases of the so-called subsidiarity principle. Notwithstanding this division of power, the EU digital agenda does explicitly address the use of AI in law enforcement and criminal proceedings. The European Commission's line of action is worth mentioning, as even when designed for internal market regulation and based on product safety, data protection – the link with fundamental rights including judicial rights – impacts the use of AI systems in criminal procedure in many ways.<sup>20</sup> Further regulatory key components are expected with the EU AI Act, adopted on 13 March 2024, which lays down harmonised rules on AI.<sup>21</sup>

However, at this point in time it is difficult to clearly forecast all the implications as, for instance, the EU AI Act's explanatory notes have not been published, and even when they are made known, manifold issues arising from the law will have to be settled. It is to be expected that these issues give rise to preliminary references where national courts will seek the guidance of the Court of Justice of the European Union (CJEU) for more clarity. Such issues will arise over time, for instance where the regulation prohibits the employment of AI in certain ways, such as various uses of 'real-time' remote biometric identification systems for law enforcement (Article 5(1)(h) EU AI Act). It classifies several AI systems relevant for law enforcement and criminal investigation as high risk, such as polygraphs or AI systems used for profiling or other risk assessments of a natural person. Of high risk are also those AI systems intended to assist judicial authorities in researching and interpreting facts and the law or in applying the law and AI systems used for the evaluation of the reliability of evidence in the course of the investigation or prosecution of criminal offences.<sup>22</sup> Using high-risk AI systems require risk management, data governance rules ensuring the quality and relevance of datasets used, technical documentation and record-keeping, transparency and the publication of information, human oversight, safeguarding the robustness of the system, and compliance with cybersecurity requirements.

A specific EU initiative to enhance criminal proceedings, in particular based on the principle of mutual recognition,<sup>23</sup> includes Regulation (EU) 2023/1543 on European Production Orders and European Preservation Orders<sup>24</sup> and Directive (EU) 2023/1544 of 12 July 2023. This legislation lays down harmonised rules for the purpose of gathering

<sup>19</sup> R Justo-Hanani, 'The Politics of Artificial Intelligence Regulation and Governance Reform in the European Union' (2022) 55 *Society of Policy Sciences* 137, available at [link.springer.com/article/10.1007/s11077-022-09452-8](https://link.springer.com/article/10.1007/s11077-022-09452-8).

<sup>20</sup> A Završnik, 'Criminal Justice, Artificial Intelligence Systems, and Human Rights' (2020) 20 *ERA Forum* 567, 576, available at [doi.org/10.1007/s12027-020-00602-0](https://doi.org/10.1007/s12027-020-00602-0).

<sup>21</sup> EU AI Act.

<sup>22</sup> Annex III EU AI Act.

<sup>23</sup> Tosza (n 14) 162.

<sup>24</sup> Regulation (EU) 2023/1543 of the European Parliament and of the Council of 12 July 2023 on European Production Orders and European Preservation Orders for electronic evidence in criminal proceedings and for the execution of custodial sentences following criminal proceedings [2023] OJ L 191/118.

electronic evidence in criminal proceedings.<sup>25</sup> It does not, however, specifically target the use of AI systems, but rather all kinds of e-evidence.<sup>26</sup>

## B. Privacy Protection as a Primary Grid for Trustworthy AI?

Data protection is often what comes to mind first when thinking about a European approach to the digital age. This has until recently been an essential part of European policy-making. But one should not forget the ECHR, even though it does not on first glance regulate AI. However, AI can infringe upon human rights and thus will end up at the Strasbourg Court sooner or later.

Europe has a long tradition of robust data protection, as exemplified by the CoE's guaranty of privacy in Article 8 ECHR; the EU set a worldwide gold standard when its GDPR took effect in 2016. The common goal in a nutshell is to ensure 'legality' and 'fairness' with data that is accurate, complete, reliable and up to date.<sup>27</sup>

### *i. Article 8 ECHR and ECtHR Case Law*

In order for the collection and processing of personal data to be justified and compliant with Article 8 ECHR, there must be clear and detailed rules concerning the scope and application of these measures. These rules should be accompanied by safeguards preventing the abuse and arbitrary use of data, for example by limiting the duration of storage, the usage, or the access by third parties of the data in question.<sup>28</sup>

The requirement for such safeguards is heightened whenever personal data is processed in an automatic or technologically sophisticated manner or is used for police purposes;<sup>29</sup> evidence-generating AI systems are used in all these instances. As such their use is to be measured against a particularly strict standard. The potential benefit of new techniques must be carefully balanced against private life interests, taking into account the as-yet unforeseeable ways that technological advancement may affect the right to private life.<sup>30</sup>

The need to justify interference with Article 8 ECHR increases with the sensitivity of the processed data. As biometric data in the form of facial images that allow for the unique identification of a person counts as sensitive data,<sup>31</sup> the use of facial

<sup>25</sup> Directive (EU) 2023/1544 of the European Parliament and of the Council of 12 July 2023 laying down harmonised rules on the designation of designated establishments and the appointment of legal representatives for the purpose of gathering electronic evidence in criminal proceedings [2023] OJ L 191/181.

<sup>26</sup> M Maras and A Alexandrou, 'Determining Authenticity of Video Evidence in the Age of Artificial Intelligence and in the Wake of Deepfake Videos' (2019) 23 *The International Journal of Evidence & Proof* 255, 256–60; Gless, Lederer and Weigend (n 6) 2–3.

<sup>27</sup> cf P De Hert, 'Data Protection as Bundles of Principles, General Rights, Concrete Subjective Rights and Rules: Piercing the Veil of Stability Surrounding the Principles of Data Protection' (2017) 3 *European Data Protection Law Review* 160.

<sup>28</sup> *S and Marper v UK* App nos 30562/04 and 30566/04 (ECtHR, 4 December 2008), para 67; *PN v Germany* App no 74440/17 (ECtHR, 11 June 2020), para 62.

<sup>29</sup> *S and Marper v UK*, para 102.

<sup>30</sup> *ibid* para 112; *Gaughran v UK* App no 45245/15 (ECtHR, 13 February 2020), para 70.

<sup>31</sup> Article 4(14) GDPR; Consultative Committee of the Convention for the protection of Individuals with regard to automatic processing of personal data (Convention 108), *Guidelines on Facial Recognition* (Council

recognition technology (FRT), especially live FRT, represents a particularly intrusive measure, requiring a high level of justification before use.<sup>32</sup> This criterion is not met where national law allows for the broad use of live FRT even for minor offences without noteworthy limitations to its use.<sup>33</sup>

In its judgment on FRT the ECtHR further referenced the Guidelines on Facial Recognition (2021), which may aid in the interpretation of the limitations for FRT set by Article 8 ECHR. In it, amongst other things, it is specified that digital images initially captured for a different purpose (eg, on social media or by surveillance cameras) are not to be used as training data for the extraction of biometric templates without an adequate legal basis for the new use, and that it is not permissible to assume the consent of the data subjects solely based on them manifestly producing or publishing the image.<sup>34</sup>

Said legal basis should, according to the guidelines, address a detailed explanation of the specific use and purpose, the minimum reliability and accuracy of the algorithm, the retention duration of the photos, the possibility of auditing these criteria, the traceability of the process, and existing safeguards.<sup>35</sup>

## *ii. EU Policy and GDPR*

The GDPR was adopted in 2016 and, among other things, requires compliance with its essential principles: lawfulness, fairness, transparency, purpose limitation, data minimisation, and accuracy. The significance of these principles for AI evidence is carved out in chapter three of this volume.<sup>36</sup> Since entering into force in 2018, the GDPR has served as a cornerstone of the EU's privacy policy with implications far beyond the originally targeted areas with regard to content and/or geography.<sup>37</sup>

Even though at first blush it may look like the GDPR falls short when it comes to adequately addressing the new challenges posed by AI evidence – as the GDPR's apparent focus is on privacy protection, not on reliability or fairness of fact-finding in criminal proceedings<sup>38</sup> – chapter three<sup>39</sup> shows that the principle-oriented law does indeed provide building blocks for trustworthy AI that can be put to use regarding the robustness of AI evidence,<sup>40</sup> with one example being the principle of data accuracy which is also enshrined in Article 4 Directive (EU) 2016/680 of the European Parliament and

of Europe, 2021) 9, available at [edoc.coe.int/en/artificial-intelligence/9753-guidelines-on-facial-recognition.html](https://edoc.coe.int/en/artificial-intelligence/9753-guidelines-on-facial-recognition.html).

<sup>32</sup> *Glukhin v Russia* App no 11519/20 (ECtHR, 4 July 2023), para 86.

<sup>33</sup> *ibid* paras 83 and 89.

<sup>34</sup> Consultative Committee of the Convention for the protection of individuals with regard to automatic processing of personal data (n 31) 9.

<sup>35</sup> *ibid* 7.

<sup>36</sup> Ch 3.V.

<sup>37</sup> ML Jones and ME Kaminski, 'An American's Guide to the GDPR' (2021) 98 *Denver Law Review* 93; C Ryngaert and M Taylor, 'The GDPR as Global Data Protection Regulation?' (2020) 114 *American Journal for International Law* 5.

<sup>38</sup> For a farsighted analysis see: M Veale et al, 'When Data Protection by Design and Data Subject Rights Clash' (2018) 8 *International Data Privacy Law* 105, 105 ff.

<sup>39</sup> Ch 3.V.

<sup>40</sup> NA Smuha, 'The EU Approach to Ethics Guidelines for Trustworthy Artificial Intelligence' (2019) 20 *Computer Law Review International* 97, 100.

of the Council of 27 April 2016 (LED),<sup>41</sup> with a specific focus on law enforcement and criminal justice.<sup>42</sup>

Yet, the promise of the GDPR's forward-oriented understanding of human agency and oversight, technical robustness and safety, data fairness, and data-quality management can only help to fend off risks posed by AI generated evidence<sup>43</sup> if the relevant stakeholders align the rather rigid legal requirements of the GDPR to issues of AI-generated evidence. For instance, with regard to the meaning of 'personal data' when data is used to generate AI evidence or when it comes to balance privacy with the free flow of data and other (business) interests: such issues arise when businesses scrape webpages for AI systems that eventually generate evidence (such as facial recognition systems) and invoke business secrecy when subpoenaed for details. In order to understand the manifold issues when invoking GDPR it is important to bear in mind that in criminal investigations and fact-finding in criminal trials, defendants and witnesses traditionally enjoy no, or little, privacy.

A promising way forward is a genuine approach to what signifies reliability for AI evidence based on GDPR's efforts for technical robustness and data quality management. It can assist new endeavours for a novel taxonomy as a basis for an assessment of reliability of AI evidence for fact-finding.<sup>44</sup> This connects, for instance, to other discussions related to Article 5 GDPR's ambition for data fairness (in particular requiring accuracy and timeliness of personal data), and transparency of data processing in relation to the individuals affected. This issue resonates with other prominent EU legislation.

The GDPR grants the right to individuals affected by profiling not be subject to a decision based solely on automated processing according to Article 22 GDPR.<sup>45</sup> Furthermore, the entitlements stipulated in Articles 13–15 GDPR specify various individual rights regarding access to information collected about them, and – in order to exercise these rights – entitlements to be notified about the data collected and to receive meaningful information about the logic involved in automated decisions.<sup>46</sup> Together, these rights build towards what is now called the right to explanation, which is at the heart of the European Agenda aiming at the protection of individuals, even though the efficiency is controversial.<sup>47</sup>

<sup>41</sup> Directive (EU) 2016/680 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and on the free movement of such data, and repealing Council Framework Decision 2008/977/JHA [2016] OJ L 119/89.

<sup>42</sup> For more details, see ch 3 and for a broader picture on the relation between data accuracy and data quality, see D Dimitrova, 'The rise of the personal data quality principle: is it legal and does it have an impact on the right to rectification?' (2021) 12 *European Journal of Law and Technology*, available at [www.ejlt.org/index.php/ejlt/article/view/768](http://www.ejlt.org/index.php/ejlt/article/view/768).

<sup>43</sup> cf Articles 5, 25 and 22 GDPR.

<sup>44</sup> E Silverman et al, 'Robot Testimony? A Taxonomy and Standardized Approach to Evaluative Data in Criminal Proceedings' in S Gless and H Whale-Bridge (eds), *Human-Robot-Interactions: A Digital Shift in Law and Its Narratives* (Cambridge University Press, 2024) 167–92.

<sup>45</sup> S Wachter et al, 'Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation' (2017) 7 *International Data Privacy Law* 76; L Edwards and M Veale, 'Slave to the Algorithm? Why a "Right to an Explanation" Is Probably Not the Remedy You Are Looking For' (2017) *Duke Law & Technology Review* 16, 18.

<sup>46</sup> A Selbst and J Powles, 'Meaningful Information and the Right to Explanation' (2017) 7 *International Data Privacy Law* 233.

<sup>47</sup> Edwards and Veale (n 45) 44 ff.

However, laws aiming for privacy protection, whilst very valuable to developing an understanding of data fairness, may only offer limited protection, for instance if these systems fall outside the scope of conventional data protection laws as law enforcement and criminal justice fall under specific legislation.<sup>48</sup> Or, in cases where data is harvested from AI systems used in everyday life (ie, devices whose primary function is assistance with daily chores, which then succumb to what has been called ‘function creep’),<sup>49</sup> it is questionable whether privacy protection is an adequate tool, or whether it would be more appropriate to employ classic defence rights. For example, the recording of a drowsiness alert, like other data stored by a vehicle,<sup>50</sup> could be a valuable source of evidence for fact-finding in criminal proceedings, especially if they concern a driver’s non-response to alerts issued by a lane-keeping assistant or drowsiness-detection system.<sup>51</sup>

It remains an unresolved issue as to how a defendant would defend against such incriminating evidence, however.<sup>52</sup> When AI-generated data produced as a result of a robot assessing human performance is proffered as evidence, a new taxonomy and a common language shared by the trier of facts and experts are required. Rules have been established for proving that a driver was speeding or intoxicated, but not for explaining the process that leads to an alert indicating the drowsiness of a human driver. These issues highlight the challenges and possibilities accompanying digital evidence, which must now be dealt with in all legal proceedings because most information is stored in electronic, rather than analogue, form.<sup>53</sup> It is welcome that supranational initiatives, such as the CoE’s Electronic Evidence Guide,<sup>54</sup> provide standards for digital evidence, although they do not take up the specific problems of evidence generated through human–robot interactions. To support the meaningful vetting of AI-generated evidence, we need a new taxonomy that distinguishes between raw, processed and evaluative data. This taxonomy will help courts find new ways to assess and test robot testimony in a reliable and fair way.<sup>55</sup>

Part of the challenge of vetting AI-generated evidence<sup>56</sup> is to support the effective use of defence rights to challenge such evidence.<sup>57</sup> It is very difficult for any fact-finder or defendant to pierce the veil of data, given that AI systems are, to a certain degree, opaque.<sup>58</sup>

<sup>48</sup> LED (n 41).

<sup>49</sup> Referring to the gradual widening of the use of a technology or system beyond the use for which it was originally intended, often: PW Grimm et al, ‘Artificial Intelligence as Evidence’ (2021) 19 *Northwestern Journal of Technology and Intellectual Property* 9, 51 f.

<sup>50</sup> Gless, Lederer and Weigend (n 6).

<sup>51</sup> S Gless et al, ‘Ca(r)veat Emptor: Crowdsourcing Data to Challenge the Testimony of In-Car Technology’ (2022) 62 *Jurimetrics* 285, 290; Gless (n 8) 195, 213.

<sup>52</sup> Gless, Lederer and Weigend (n 6) 29.

<sup>53</sup> PW Grimm, DJ Capra and GP Joseph, ‘Authenticating Digital Evidence’ (2017) 69 *Baylor Law Review* 1, 1f.

<sup>54</sup> ‘IPROCEEDS-2: Launching of the Electronic Evidence Guide v30’ (Council of Europe, 22–23 June 2022), available at [www.coe.int/en/web/cybercrime/-/iproceeds-2-launching-of-the-electronic-evidence-guide-v-3-0#](http://www.coe.int/en/web/cybercrime/-/iproceeds-2-launching-of-the-electronic-evidence-guide-v-3-0#).

<sup>55</sup> Silverman et al (n 44).

<sup>56</sup> N Ommen et al, ‘Toward a Better Understanding of Stakeholder Participation in the Service Innovation Process: More than One Path to Success’ (2016) 69 *Journal of Business Research* 2409, 2409.

<sup>57</sup> Gless (n 8) 195, 232–50.

<sup>58</sup> C Rudin, ‘Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead’ (2019) 1 *Nature Machine Intelligence* 206, 206, available at [www.nature.com/articles/s42256-019-0048-x](http://www.nature.com/articles/s42256-019-0048-x).

## C. EU Agenda on Obtaining and Regulating Novel Forms of Evidence

Europe has a tradition of regulating novel forms of evidence in a principle-based way. The Budapest Convention (see above I.A.i) illustrates this prominently for the work in the CoE forum. The EU e-evidence package (see above I.A.ii) illustrates the impact of the line from Brussels on the European legal landscape, despite the so-called subsidiarity principle that prioritises the competence of Member States in criminal justice (see above I.A.ii). These legislative initiatives, however, are in many ways a continuation of traditional harmonisation and cooperation efforts. Under the e-evidence package, judicial authorities will be able to directly request electronic evidence from service providers based in another Member State, and service providers will be obliged to respond swiftly.<sup>59</sup>

A new chapter has been opened with the prospects of the use of AI systems, with their specific capability of replacing human decision-making and action, or even surpassing human capabilities. This impact on fact-finding is illustrated by the explanations provided in different chapters in this volume. As explained, AI employment offers both promise and possible dangers in obtaining and assessing evidence. The EU AI Act acknowledges this by explicitly stating that

AI systems intended to be used by or on behalf of law enforcement authorities, or by Union institutions, bodies, offices or agencies, in support of law enforcement authorities to evaluate the reliability of evidence in the course of the investigation or prosecution of criminal offences

are to be classed as high risk.<sup>60</sup> Furthermore, the EU AI Act goes hand-in-hand with the EU e-evidence package when relying on a public–private partnership.<sup>61</sup>

The EU AI Act, composed of a risk-based approach to product safety and fundamental rights protection, does not explain the legal ramifications of such a blend: If an AI system is classified as ‘high risk’ (as opposed to being labelled an ‘unacceptable risk’ or only a ‘limited risk’ or a ‘minimal risk’), the EU AI Act imposes a strict requirement on the responsible producer and deployer (often the person with responsibility of ensuring the safe and compliant use of AI systems when they are rolled out as opposed to the end user) for transparency, documentation and monitoring.

Nevertheless, as there is no explicit reference to the use of AI for criminal justice (beyond the expressly named high-risk system in Article 6 and Annex III, including polygraphs or facial recognition) many questions remain regarding the EU AI Act’s impact on evidentiary issues. The law generally aims to safeguard not only product safety, but also fundamental rights and the rule of law by containing risks linked to the employment of AI. In fact, this ground-breaking act also regulates evidentiary issues, for instance when it prohibits biometric categorisation systems that use sensitive

<sup>59</sup> Tosza (n 14) 161, 170.

<sup>60</sup> EU AI Act Annex III(6)(c).

<sup>61</sup> S Tosza, ‘The E-Evidence Package is Adopted: End of a Saga or Beginning of a New One?’ (2023) 2 *European Data Protection Law Review* 163, 164.

characteristics or untargeted scraping of facial images from the internet or CCTV footage to create facial recognition databases. Yet, apart from regulation of strikingly controversial uses of AI, the act offers little guidance on how to handle AI-generated evidence in other situations (for more details see below II.C.iii), leaving it to the respective authorities and courts to find their way in handling the specific risks of the manifold ways to use AI systems for evidentiary purposes.

Such self-restraint on evidentiary questions complies with the macro-managing EU approach to AI that does not regulate distinctive AI systems, but requires users to understand the risks of using an AI system in the relevant situation. It also aligns with the fact that the admission and evaluation of evidence is a domestic matter (see below II.C.ii), which often remains under-regulated and left to case law.<sup>62</sup> Given the emphasis on a novel approach to regulation and European-wide harmonisation as well as on a human-centric approach, one would have expected more clear red lines.<sup>63</sup> It is telling that law enforcement is regulated under a separate set of rules, allowing the use of biometric identification systems, albeit only within a narrow, exceptional circumstance: in publicly accessible spaces subject to prior judicial authorisation and for a strictly defined list of crimes.<sup>64</sup> ‘Real-time’ biometric identification must comply with strict requirements and is limited to a few specific uses, such as targeted searches for victims (abduction, trafficking, sexual exploitation) or prevention of a specific and present terrorist threat.

## II. A European Perspective on AI and AI-Generated Evidence

How does the European approach play out in cases of AI employment in criminal justice, in particular when AI is employed to obtain evidence or when AI-generated evidence is used for fact-finding in a criminal trial? The focus of this section is on evidence autonomously generated by AI systems, thus by using some form of machine learning, sophisticated mathematical methods, and big data processing. AI use in law enforcement, such as predictive policing, or AI use in general judicial decision-making, such as for pre-trial detention, sentencing or corrections, will be explained briefly where necessary for a better understanding of the broader context.

<sup>62</sup> L Macula, ‘The Potential to Secure a Fair Trial Through Evidence Exclusion: A Swiss Perspective’ in S Gless and T Richter (eds), *Do Exclusionary Rules Ensure a Fair Trial? A Comparative Perspective on Evidentiary Rules* (Springer Open, 2019) 15, 40 ff; T Weigend, ‘The Potential to Secure a Fair Trial Through Evidence Exclusion: A German Perspective’ in S Gless and T Richter (eds), *Do Exclusionary Rules Ensure a Fair Trial? A Comparative Perspective on Evidentiary Rules* (Springer Open, 2019) 61, 87 f.

<sup>63</sup> M Veale and F Zuiderveen Borgesius, ‘Demystifying the Draft EU Artificial Intelligence Act’ (2021) 22 *Computer Law Review International* 97, 112.

<sup>64</sup> I Barkane, ‘Questioning the EU Proposal for an Artificial Intelligence Act: The Need for Prohibitions and a Stricter Approach to Biometric Surveillance’ (2022) 27 *Information Polity* 147, available at [content.iospress.com/download/information-polity/ip211524?id=information-polity%2Fip211524](https://content.iospress.com/download/information-polity/ip211524?id=information-polity%2Fip211524).

## A. The Strasbourg Imprint on Criminal Justice: A Right to a Fair Trial

Overall, the European perspective is an overarching approach rather than one aimed at evidentiary issues specifically, given that criminal proceedings are the domain of the territorial states. However, criminal justice systems have increasingly been shaped by European regulations over the last few years. One example is the mark the ECHR has left on domestic criminal justice issues; also illustrative are the various topics taken up by CDPC, such as harmonisation rules for mutual legal assistance in criminal matters. Furthermore, the EU institutions not only changed the European law enforcement and criminal justice landscape with Europol, a European Arrest Warrant and a European Prosecutor, but also furnished the larger picture when aiming for trustworthy AI and the strengthening of human oversight.

Yet, for AI-generated evidence human rights might be decisive in criminal justice. Even if the ECHR does not stipulate specific protection against AI-generated evidence, general rules, such as the right to examine witnesses against oneself (Article 6(3)(d) ECHR) can help establish the overall fairness of criminal trials, including equality of arms. Before exploring these rules, it is important to keep in mind that the case-by-case approach and deference to the national legislature and administrative bodies leaves a wide margin of discretion as how to protect human rights in their domestic system.<sup>65</sup> Thus, one cannot draw on a comprehensive, strict standard regarding evidentiary issues based on ECtHR case law.<sup>66</sup> But one can use the skeleton of relevant rules developed in the continuous jurisprudence of the Strasbourg Court to explore possible European limits to the use of AI systems for evidentiary purposes.

## B. Using AI to Detect Crime

Criminal proceedings start with the suspicion that a crime has occurred, and possibly that a specific person has committed it. Given the focus on legal ramifications of AI-generated evidence, it is important to address the (blurring) line between AI-generated leads and AI evidence. There is a difference between the generation of mere leads by AI systems, the assistance of such systems in the assessment of information, and AI-driven generation of 'hard evidence'.<sup>67</sup>

From a legal point of view, the emergence of a suspicion is crucial for many reasons. Under certain circumstances, the mere suspicion that a crime has been committed can have consequences in the administrative realm, as has been illustrated by the so-called Dutch daycare benefit data scandal.<sup>68</sup> The Dutch tax authorities first set out merely to create risk profiles in their effort to spot childcare benefit fraud, and they penalised the

<sup>65</sup> Gless and Martin (n 9) 36.

<sup>66</sup> K Ligeti et al, 'Admissibility of Evidence in Criminal Proceedings in the EU' *eu crim* (2020) 201, 205–06.

<sup>67</sup> S Gless, 'Cross-border Admissibility of AI-Evidence' in Lelieur (n 7) 359. See also ch 1.II and III.

<sup>68</sup> R Peeters et al, 'Administrative exclusion in the infrastructure-level bureaucracy: The case of the Dutch daycare benefit scandal' (2023) *Public Administration Review* 863, 877.

blacklisted families based on the system's risk indicators, until it turned out that the authorities lacked a statutory basis for such processing of personal data.

Even more important from the point of view of this project is the relevance of such suspicion as a justification for the government to use intrusive measures characteristic of a criminal investigation, such as house searches, DNA sample tests or arrests. In this area Forensic AI systems play a prominent role, as explained in the chapters with examples of filtering AI systems (eg, Hansken); data-mining AI systems and facial recognition techniques (eg, NeoFace Watch, Clearview); AI voiceprint systems or probabilistic genotyping AI systems (eg, TrueAllele, STRMix).<sup>69</sup>

But such suspicion can also be substantiated by data generated by Consumer Product AI used in everyday life, such as GPS or tracking data (eg, resulting from the use of Google Earth, Find My iPhone, Alexa, etc).<sup>70</sup> In particular, the harvesting of data from Consumer Product AI creates new challenges to the traditional understanding of the significance of suspicion for legal proceedings. Analysis of data by AI can be used to generate suspicion, for instance with natural language-based analysis of tax documents,<sup>71</sup> retrospective analysis of GPS locations stored in smartphones,<sup>72</sup> or even more vague data-profiling of certain groups.<sup>73</sup> In each of these cases, AI systems create a suspicion that allows the authorities to investigate and possibly prosecute a crime – one that would not have come to the government's attention previously.<sup>74</sup>

Today, surveillance systems are the most controversially debated example, as they raise issues of privacy, over-policing, and potentially discrimination and creating risks for democracy. The EU AI Act's prohibition of biometric identification systems in publicly accessible spaces (see above I.C) is one prominent example.

Broader criminal justice issues connected to these AI systems arise from the complex methods used to build these tools, including machine learning methods. Both the design and the data used for it might be skewed. The system may be built to reflect and perpetuate structural prejudices when using data sets from previous prosecution work as training material. If the risk of bias is not addressed (and countered), human bias, already present in the criminal justice system, can be reinforced by biased training data, insufficiently calibrated machine learning, or both. This can result in ineffective predictive tools that either do not identify 'true positives', that is, the people at risk of committing a crime, or that burden the public or a specific minority with unfair and expensive over-policing.<sup>75</sup> In any case, a risk assessment is a prognosis, and as such it

<sup>69</sup> For a definition and a more detailed analysis on Forensic AI evidence, see ch 1.II ff.

<sup>70</sup> For a definition and a more detailed analysis on Consumer Product AI evidence, see ch 1.II ff.

<sup>71</sup> A Calafato et al, 'A Controlled Natural Language for Tax Fraud Detection' in B Davis et al (eds), *Controlled Natural Language* (Springer, 2016) 1, 3 ff.

<sup>72</sup> J Moore et al, 'Find Me If You Can: Mobile GPS Mapping Applications Forensic Analysis & SNAV the Open Source, Modular, Extensible Parser' (2017) 12 *Journal of Digital Forensics, Security and Law* 15, 25.

<sup>73</sup> K Kremens and WJ Jasiński, 'Editorial of Dossier "Admissibility of Evidence in Criminal Process. Between the Establishment of the Truth, Human Rights and the Efficiency of Proceedings"' (2021) 7 *Revista Brasileira de Direito Processual Penal* 15, 31.

<sup>74</sup> M Hildebrandt, *Smart Technologies and the End(s) of Law: Novel Entanglements of Law and Technology* (Edward Elgar, 2015) 159–85; M Zaia, 'Forecasting Crime? Algorithmic Prediction and the Doctrine of Police Entrapment' (2020) 18 *Canadian Journal of Law and Technology* 255, 262.

<sup>75</sup> AG Ferguson, 'Policing Predictive Policing' (2016) 94 *Washington University Law Review* 1109; S Gless, 'Predictive Policing – In Defense of "True Positives"' in E Bayamlioglu, I Baraliuc, LAW Janssens et al (eds),

always carries its own risks because it cannot be checked entirely; such risk assessments therefore raise ethical and legal issues when used as the basis for action.<sup>76</sup>

The EU AI Act took a somewhat different angle. While the GDPR focuses on data collection and data processing, the EU AI Act aims at advancing what has been coined ‘trustworthy AI’, based on a combination of product safety and fundamental rights protection. The EU strategy involves various overarching concerns, including criminal justice topics (see above I.C) and evidentiary issues (see below II.C.ii).

## C. Criminal Investigation and Fact-Finding in Criminal Proceedings

Criminal investigations and fact-finding are at the heart of criminal proceedings. Particularly in Europe, the search for truth shapes each step in a criminal trial. The opportunity to harvest and process huge pools of data and, in doing so, obtain evidence for a criminal case or learn about presumably hidden patterns of undetected crimes has changed the perception of criminal investigations and fact-finding in criminal proceedings to a certain degree.

### *i. CDPC’s Initiative on AI and Criminal Justice*

Such new opportunities raise new questions, in particular where states cooperate across borders to fight crime. The controversial *Ennetcon* and *Encrochat* cases illustrate this: information retrieved from so-called cryptophones (ie, phones that use encryption for anonymous communication) by the *Hansken* system has been accepted in certain states, but not in others (see below II.C.ii).<sup>77</sup>

Anticipating evidentiary issues connected to the use of AI, a CDPC working group outlined possibilities for a CoE instrument on AI and criminal law with a feasibility study.<sup>78</sup> It explores the potential to pave the way for the adoption of an international legal instrument that also addresses evidentiary problems. It is important to note that the CDPC working group did not restrict itself to considering the use of AI systems designed for forensic purposes, but included AI in everyday life where AI systems monitor humans and thus are in a position to provide information.<sup>79</sup> For example, a modern car records manifold data on its user, including infotainment and braking

*Being Profiled: Cogitas Ergo Sum: 10 Years of Profiling the European Citizen* (Amsterdam University Press, 2018) 76.

<sup>76</sup> M Browning and BA Arigo, ‘Stop and Risk: Policing, Data, and the Digital Age of Discrimination’ (2021) 46 *American Journal of Criminal Justice* 298, 310; OJ Gstrein et al, ‘Ethical, Legal and Social Challenges of Predictive Policing’ (2019) 3 *Católica Law Review, Direito Penal* 77, 86–88.

<sup>77</sup> Ch 1.VII and ch 7.V.D.

<sup>78</sup> CDPC, *Feasibility Study on a Future Council of Europe Instrument on Artificial Intelligence and Criminal Law* (Council of Europe, 4 September 2020), available at [rm.coe.int/cdpc-2020-3-feasibility-study-of-a-future-instrument-on-ai-and-crimina/16809f9b60](https://rm.coe.int/cdpc-2020-3-feasibility-study-of-a-future-instrument-on-ai-and-crimina/16809f9b60).

<sup>79</sup> CDPC, *Working Paper for the meeting of 27 March 2019* (Council of Europe, 5 February 2019), available at [rm.coe.int/cdpc-2019-3-working-paper-for-cdpc-working-group-of-experts-on-artific/1680925b9f](https://rm.coe.int/cdpc-2019-3-working-paper-for-cdpc-working-group-of-experts-on-artific/1680925b9f).

characteristics.<sup>80</sup> During automated driving, driving assistants such as lane-keeping assistants or drowsiness-detection systems monitor drivers to ensure they are ready to respond to a take-over request if required.<sup>81</sup> If an accident occurs, this kind of alert could be used in legal proceedings in various ways.<sup>82</sup>

With the publication of its feasibility study in 2020,<sup>83</sup> the working group aimed for common standards and to clarify connected procedural issues as well as possible human rights implications.<sup>84</sup>

The working group looked at evidentiary issues concerning the admissibility and reliability of such data, but it also looked at problems of cooperation if data must be transferred from one country to another for legal proceedings; for example, if automated vehicles cause accidents when used in other countries or if other illegal activity linked to autonomous driving affects more than one jurisdiction. The challenges of conserving and obtaining evidence in alleged cases of criminal activity connected to automated driving also shows the importance of including private sector stakeholders. Based on answers to a questionnaire sent back by governments, the working group highlighted that automated driving generates valuable data that can also be of great interest to criminal investigations. This will require new evidentiary rules, for instance with regard to accessibility and readability of data, in particular AI-generated evidence,<sup>85</sup> as traditional rules may not be designed to meaningfully test the reliability and credibility of this new form of evidence.<sup>86</sup> The committee will continue its work.

Credit must be given to the CDPC's AI and Criminal Law working group for tackling the evidentiary issue in a working paper that proposes novel regulation. This regulation concerns not only retrieving data from AI systems, but also a cross-border production order for both the prosecution service and defendants in order to give them access to sufficient information for a meaningful defence in compliance with the rights set out in Article 6 ECHR.<sup>87</sup> Furthermore, it is an obligation for domestic systems to adopt such legislative and other measures as may be necessary to test the trustworthiness of data retrieved from AI systems.<sup>88</sup>

## *ii. EU AI Act*

Whether – or rather how – the EU AI Act applies to AI generated evidence in criminal cases remains an open issue at this point. On the one hand, Article 6(2) and Annex III the EU AI Act stipulate specific rules for certain high-risk AI systems, such as polygraphs or facial recognition, explicitly designed for law enforcement and criminal justice. On the

<sup>80</sup> N Le-Khac et al, 'Smart Vehicle Forensics: Challenges and Case Study' (2020) 109 *Future Generation Computer System* 500.

<sup>81</sup> Gless et al (n 51) 286.

<sup>82</sup> Gless (n 8) 195; Gless, Lederer and Weigend (n 6) 2–3.

<sup>83</sup> CDPC (n 78).

<sup>84</sup> D Leslie et al, *Artificial intelligence, human rights, democracy, and the rule of law: a primer* (Council of Europe and the Alan Turing Institute, 2021).

<sup>85</sup> CDPC (n 2).

<sup>86</sup> Gless (n 8) 195.

<sup>87</sup> CDPC, *Drafting Committee to elaborate an instrument on Artificial Intelligence and Criminal Law (CDPC-AICL) Framework document* (Council of Europe, 28 April 2022) 1.

<sup>88</sup> *ibid.*

other hand, the EU AI Act is mute on the applicability of its general rules to fact-finding in criminal cases. This raises practical issues. For instance, if a prosecution service reads out a car's AI systems, does it adopt the role of a deployer (with certain obligations) or a mere user who must comply with the relevant documentation and monitoring rules? One could argue that in its general rules, the EU AI Act only establishes/stipulates rules for ordinary producers, deployers and users of AI systems, but not law enforcement agents, prosecution services, or courts. Yet, such an argument would ignore the broad ambition of the EU AI Act with its combination of product security and fundamental rights protection. There is no apparent reason why the general regime should not apply to AI systems' use in criminal justice. In particular, Article 2 EU AI Act stipulating the scope of the legislation does not exclude law enforcement or criminal prosecution. On the contrary, the EU AI Act's case-specific regulation for AI tools expressly designed for use in criminal justice in Annex III<sup>89</sup> includes high-risk

AI systems intended to be used by or on behalf of law enforcement authorities, or by Union institutions, bodies, offices or agencies, in support of law enforcement authorities to evaluate the reliability of evidence in the course of the investigation or prosecution of criminal offences;

and

AI systems intended to be used by a judicial authority or on their behalf to assist a judicial authority in researching and interpreting facts and the law and in applying the law to a concrete set of facts, or to be used in a similar way in alternative dispute resolution.<sup>90</sup>

When Annex III, para 6(b) explicitly names 'AI systems intended to be used by or on behalf of law enforcement authorities or by Union institutions, bodies and agencies in support of Law enforcement authorities as polygraphs and similar tools',<sup>91</sup> it does not exclude other AI systems from falling under the general rules.

It is to be expected that the issue will be settled by case law. It will have to take into account the many facets of this legislation: the blanket nature of the legal instrument, explicitly dealing with specific evidentiary issues (such as the use of polygraphs or facial recognition); the basis in Article 114 of the Treaty on the Functioning of the European Union, which mandates law-making to ensure the establishment and functioning of the internal market; and its position in the EU digital single-market strategy aiming to ensure that AI is safe, and is developed and used in compliance with fundamental rights obligations based on common rules. Partly, the EU AI Act is meant to fend off a fragmentation of the internal market that could result as a consequence of particular domestic requirements for AI systems. The regulation on evidentiary issues is rather a spillover effect, for instance when governing on Consumer Product AI. This is important because, as explained at I.B.ii above, that data is harvested from AI systems used in everyday life. Other rules clearly draw a red line on certain practices, to uphold the rule of law and fend off risks to infringements on human rights, such as the general prohibition against using AI systems for real-time biometric identification.

<sup>89</sup> Annex III EU AI Act.

<sup>90</sup> Annex III, para 8 EU AI Act.

<sup>91</sup> Annex III, para 6(b) EU AI Act.

In the long run, the legal medley of product safety and fundamental rights protection might prove not a weakness but a strength for a robust grip on AI evidence. The logic of 'product safety regulation' could develop into a basis for testing the reliability of AI evidence. A strict regime applies to AI systems from design to development and deployment regarding data governance, documentation and record-keeping, transparency and information to users. This can be also used to safeguard human oversight, robustness, accuracy and security when AI-generated evidence is used for fact-finding in criminal cases. The goal is a taxonomy for AI evidence. Such a taxonomy ought to be structured around the relevant types of data, in particular: raw data; measurement data; and evaluative data. Raw data covers unprocessed, machine-generated recordings of physical indicators (for instance, distance or temperature). When raw data is processed by algorithms one receives measurement data, which can be understood by fact-finders on condition that the processing algorithms are disclosed. For example, wheel speed sensors in a vehicle calculate speed, helping experts assess vehicle control. Evaluative data, unlike the other two data groups, is a new category of processed data that includes data generated autonomously by AI. Unlike measurement data, evaluative data cannot be fully reconstructed due to the complexity of processing, in particular where machine learning is involved. This leads to challenges in understanding and interpreting the information proffered as evidence in a criminal trial. It is crucial to interpret such data in the light of what in forensics is called 'circumstantial information', which provides context for AI evidence, such as traffic data for an assisted driving system's observations. Courts must combine all relevant data with legal propositions to determine the reliability and the relevance to a case.<sup>92</sup>

The primary focus of this chapter is on reliability of AI-generated evidence. In this respect the EU AI Act's rules for high-risk AI systems' performance throughout their lifecycle can be helpful as they require an appropriate level of accuracy, robustness and cybersecurity in accordance with the state of the art. Before the adoption of the AI Act such requirements were set only for specific areas (such as vehicles; *cf* Article 2(2) and Annex I). To determine the reliability of evidence used for fact-finding in legal proceedings it is important to understand the level of accuracy and accuracy metrics of a tool; such information should therefore be communicated to the users.<sup>93</sup> Against this background it is difficult to understand why highly regulated areas, such as type-approved motor vehicles remain outside the scope of the EU AI Act. If accuracy is to become a proxy for trustworthy AI, the logical consequence should not only be that AI systems whose scientific validity is not proven are flagged out as not trustworthy, but also that all systems that can conflict essential values of the EU – like a fair trial – should be treated as high risk for judicial rights or be prohibited (a consequence drawn, *eg*, for polygraphs, Annex III, paras 6(b) and 7(a)).

In order to safeguard procedural fundamental rights, such as the right to an effective remedy and to a fair trial as well as the right of defence and the presumption of innocence, AI systems used for evidentiary issues must be transparent, explainable and documented. One important element will be the quality of data used; the EU AI Act reflects this when naming dangers to fundamental rights: in light of the intended

<sup>92</sup> Silverman et al (n 44).

<sup>93</sup> *cf* Art 15 EU AI Act as well as Recital 49.

purpose of high-risk AI systems, the quality of data sets used, technical documentation and record-keeping, transparency and the provision of information to users, human oversight, and robustness, accuracy and cybersecurity must be ensured.<sup>94</sup>

### *iii. Admissibility and Evaluation of Evidence – Domestic Decisions*

Criminal justice is a domestic domain, even though European institutions have gained significant influence over the last decades. Accordingly, the admissibility of a piece of evidence is the decision of the domestic court or other authority in charge of a criminal proceeding.<sup>95</sup> Given the ubiquitous use of AI systems one would assume that AI-generated evidence should be easily transferable across borders and encounter minimal hurdles regarding admissibility in different jurisdictions.<sup>96</sup>

In fact, the use of AI systems in the evidentiary process might revivify the debate on the preconditions of reliable evidence. As it is based on mathematical concepts and computer science and offers new opportunities to conceptualise uncertainties attached to different modes of proving facts,<sup>97</sup> AI use could open up new ways of sharing evidence across borders without a loss of information.

Yet, with the differing procedural rules for the obtainment and assessment of evidence in each domestic system, there is an ongoing problem: information obtained under the rules of one legal system cannot expect universal applicability when presented as evidence in another different legal system. This is even true in cases where forensic experts use natural sciences (eg, biology, chemistry or physics) to assist a court in establishing a fact-based examination and analysis of evidential material to report to the court.

A particular vulnerable point when using AI systems to obtain or assess evidence could be the fair trial's maxim of equality of arms based on Article 6 ECHR, defined by the Strasbourg Court as the defendants' chance to effectively influence a criminal court's decision when standing trial. This can work in two directions: first, fair-trial questions can arise regarding the lawfulness of gathering evidence using AI system.<sup>98</sup> When the Hansken AI system was used to retrieve data for substantiating prosecutions from huge digital data sets after authorities hacked into the cryptophone systems, defence questioned, among other things, the reliability and even-handedness of the Hansken system and the evidence gathered through it. But the courts rejected the different motions, for instance arguing that the Hansken system was merely used to view (not even to gather) evidence already collected so that no specific legal basis was needed for its use.<sup>99</sup> Without going into too much detail, it nevertheless seems a valid argument that for equality of arms, the defence must have a comparable opportunity to that of the prosecution to

<sup>94</sup> cf Arts 10–15 EU AI Act as well as Recital 43.

<sup>95</sup> S Gless, 'Bird's-Eye View and Worm's-Eye View: Towards a Defendant-Based Approach in Transnational Criminal Law' (2015) 6 *Transnational Legal Theory* 117; S Gless and P Pfirter, 'Cross-Border Access and Exchange of Digital Evidence: Cloud Computing Challenges to Human Rights and the Rule of Law' in V Mitsilegas and N Vavoula (eds), *Surveillance and Privacy in the Digital Age* (Hart Publishing, 2021) 5.

<sup>96</sup> Gless (n 67) 353.

<sup>97</sup> B de Finetti, 'Bayesianism: Its Unifying Role for Both the Foundations and Applications of Statistics' (1974) 42 *International Statistical Review* 117, 121.

<sup>98</sup> Ch 7.V.D.

<sup>99</sup> District Court of Amsterdam, judgment of 19 April 2018, ECLI:NL:RBAMS:2018:2504 (case nr 13/997097-16), para 7.3.

search a dataset using the tool the prosecution has used, possibly with repeated and refined searches.<sup>100</sup> Furthermore, one can argue that the defence should receive certain information about AI systems used to understand possible blind spots and what kind of exonerating information it may not have retrieved as a consequence.

Extending this right to confront incriminating evidence beyond the analogue world into the digital era, a creative approach is needed. On the basis of decisions rendered by the ECtHR concerning large datasets and Article 6 ECHR,<sup>101</sup> one can argue that the defence is entitled to meaningful access to all relevant data.<sup>102</sup> The aim for the digital era is to develop an option to meaningfully challenge the reliability of AI-generated evidence. Such a path could also be opened based on the right to examine an incriminating witness under Article 6(3)(d) ECHR, seeking to achieve ‘knowledge parity’ and meet the benchmark of European case law on Article 6 ECHR and its notion of procedural fairness.<sup>103</sup> In cases in which the court cannot grant sufficient access to information for defendants and cannot provide sufficient validation for reliability of an automated generation of evidence in other ways, it must exclude the evidence from the adjudication on the defendant’s guilt in order not to violate the basic right to a fair trial.<sup>104</sup>

If an AI system is used either to obtain or to assess evidence incriminating an individual, it is questionable whether the defendant will be able to contest this assessment effectively unless they have access to information about the design of the system, its training data, the mathematical methods employed, etc.<sup>105</sup>

The chapters in this volume illustrate that AI-generated evidence, like any other evidence, must be reliable, valid and credible to be admitted in trial. The EU package on e-evidence, like previous regulation modifying the transfer of evidence – based on the principle of mutual recognition – is not matched by changes on the domestic level. Questions regarding the use of such evidence, with regard to the reliability and fairness of fact-finding, remain a purely domestic affair.<sup>106</sup> The competent authorities in the respective jurisdiction in charge of deciding on the merits of a case must decide on the admissibility of a certain piece of evidence.<sup>107</sup> The underpinning rationale lies in the differences of requirements that guarantee reliability and fairness in each jurisdiction, which are core requirements for the acceptance of a judgment.<sup>108</sup> Despite the development in forensics, the ‘legal connotation’<sup>109</sup> still governs admissibility.

<sup>100</sup> B Custers and L Stevens, ‘The Use of Data as Evidence in Dutch Criminal Courts’ (2021) 29 *European Journal of Crime, Criminal Law and Criminal Justice* 25, 40. See also District Court of Amsterdam (n 99).

<sup>101</sup> *Sigurður Einarsson and Others v Iceland* App no 39757/15 (ECtHR, 4 June 2019); *Rook v Germany* App no 1586/15 (ECtHR, 25 July 2019).

<sup>102</sup> M Galič, ‘De rechten van de verdediging in de context van omvangrijke datasets en geavanceerde zoekmachines in strafzaken: een suggestie voor uitbreiding’ (2021) 2 *Boom Strafblad* 41; Gless, Lederer and Weigend (n 6) 1.

<sup>103</sup> J Jackson and S Summers, *The Internalisation of Criminal Evidence: Beyond the Common Law and Civil Law Traditions* (Cambridge University Press, 2012) 69–70, 79–80.

<sup>104</sup> Gless (n 8) 195.

<sup>105</sup> A Sachoulidou, ‘Going beyond the “Common Suspects”: To Be Presumed Innocent in the Era of Algorithms, Big Data and Artificial Intelligence’ (2023) *Artificial Intelligence and Law*.

<sup>106</sup> cf M Gialuz and S Quattrocchio, ‘Predictive Justice in Italy’ in Lelieur (n 7) 195, 201.

<sup>107</sup> S Gless, *Internationales Strafrecht: Grundriss für Studium und Praxis*, 3rd edn (Lichtenhahn, 2021) no 267, 97.

<sup>108</sup> Jackson and Summers (n 103) 69–70.

<sup>109</sup> S Gless, ‘Grenzüberschreitende Beweissammlung’ (2013) 125 *Zeitschrift für die gesamte Strafrechtswissenschaft* 573.

The question is whether, in Europe, one can find a way to develop a standard on how to define reliability and, for instance, handle possible sources of error in Forensic AI evidence by assigning a proper evidentiary value on a DNA sample match or a GPS location, taking into account inherent methodological error rates, etc.<sup>110</sup> These questions have been central in a debate that has been fuelled by forensic failure, notably the Danish scandal in which GPS location data turned out to be erroneous.<sup>111</sup>

While the reluctance to blindly accept evidence from other countries as reliable in one's own system can be understood as the heritage of a European criminal justice landscape where different legal systems led to divergent national rules on admissibility and exclusion of evidence in an environment of free evaluation of evidence,<sup>112</sup> it is noteworthy that this volume finds a general tendency to admit AI-generated evidence without too much scrutiny of their validity, reliability or credibility. Yet, just as with other types of evidence, it might move from being blindly trusted to being under increased scrutiny. Maybe AI-generated evidence will change the traditional understanding in inquisitorial systems that the trial judge or the investigating judge may base their belief solely on their appreciation of a certain piece of evidence that seems most reliable in the eyes of the court.

To meaningfully challenge the reliability (and consequently the admissibility) of AI output, the defence needs to demonstrate that it is either not valid and/or not reliable and therefore needs access not only to the case file but to the raw data,<sup>113</sup> specifications regarding the data processing, such as the AI's source code, its original specifications, its intended purpose and its training dataset.<sup>114</sup> Such access to data raises many issues, including how to enforce such a right in a public–private partnership when a company invokes a propriety entitlement, such as business secrets.<sup>115</sup> The EU AI Act does not address these questions, nor does any other EU instrument reflect on the manifold practical issues.

## D. Institutional Safeguards and Defence Rights

The employment of AI systems will change the situation for criminal defence profoundly.<sup>116</sup>

The use of AI systems in law enforcement and criminal investigations, and the omnipresence of AI devices that monitor the daily life of humans, impact the criminal

<sup>110</sup> J Vuille and F Taroni, 'Measuring Uncertainty in Forensic Science' (2021) 24 *IEEE Instrumentation & Measurement Magazine* 5.

<sup>111</sup> L Wachter Lentz and N Sunde, 'The Use of Historical Call Data Records as Evidence in the Criminal Justice System—Lessons Learned from the Danish Telecom Scandal' (2021) 18 *Digital Evidence & Electronic Signature Law Review* 1, 3 ff.

<sup>112</sup> cf T Weigend, 'The Potential to Secure a Fair Trial Through Evidence Exclusion: A German Perspective' in S Gless and T Richter (eds), *Do Exclusionary Rules Ensure a Fair Trial? A Comparative Perspective on Evidentiary Rules* (Springer Open, 2019) 61, 62 and 72 ff.

<sup>113</sup> For more detailed explanations on the significance of different data types, see Silverman et al (n 44).

<sup>114</sup> Gless, Lederer and Weigend (n 6).

<sup>115</sup> R Wexler, 'Life, Liberty, and Trade Secrets: Intellectual Property in the Criminal Justice System' (2018) 70 *Stanford Law Review* 1343, 1413 ff.

<sup>116</sup> Gless (n 8) 195, 199 f.

trial,<sup>117</sup> as fact-finding might shift from traditional human-centric investigations to an approach dedicated to data analysis.

This shift could erode the effectiveness of traditional defence rights.<sup>118</sup>

The design of an AI system, the choice of training methods, and incorrect or incomplete data can all lead to limited reliability of evidence generated by that system. This raises ethical and legal questions with regard to procedural justice (for instance, with regard to the right to a fair trial). Since AI is complex and its workings can be non-transparent or hard to explain, it may be difficult for suspects to defend themselves against this. If decisions in criminal law procedures increasingly rely on the results of AI, this could lead to situations similar to those in Kafka's novel *The Trial*, in which suspects do not know what they are accused of, where the accusations come from, and on which information (data, analysis, conclusions) these accusations are based.

In the future, meaningful defence can be achieved through different paths: the ECtHR could fine-tune its understanding of a fair trial and equality of arms or a right to confrontation with novel judgments explaining what the digital turns mean under Article 6 ECHR.<sup>119</sup> Or the CoE could adopt a new policy in the CDPC forum, establishing cross-border standards for obtaining and admitting evidence.<sup>120</sup>

The relevant authorities or courts, nudged by litigation, could also put into detail how Article 5 GDPR can be used to safeguard reliability of evidence by ensuring data quality or how Articles 22, 13, 14 and 15 GDPR and its right to explanation can be used to enhance accountability and transparency, as humans must achieve sufficient knowledge of how the machine generates its outputs.<sup>121</sup> If such a right to an explanation is interpreted functionally, it will help criminal defence.

It remains to be seen whether future courts and legal scholarship will be able to provide a new understanding of basic principles in criminal proceedings, such as the presumption of innocence. A new understanding is needed in view of the possibility that investigative powers will be exercised on individuals who are not the subjects of criminal investigations but instead predictive policing,<sup>122</sup> as these individuals would not be offered traditional procedural protections. This is a complex issue doctrinally, because in Europe the presumption of innocence only applies after the charge. If there is no charge, there is, in principle, no protection. However, once a charge is levelled, the protection applies retroactively.

<sup>117</sup> M Hildebrandt and BJ Koops, 'The Challenges of Ambient Law and Legal Protection in the Profiling Era' (2010) 73 *Modern Law Review* 428, 437 f.

<sup>118</sup> H Oran, 'Does Brady Have Byte: Adapting Constitutional Disclosure for the Digital Age' (2016) 50 *Columbia Journal of Law & Social Problems* 97.

<sup>119</sup> Gless (n 8) 195; S Gless and T Weigend, 'Intelligente Agenten als Zeugen im Strafverfahren?' (2021) 76 *Juristenzeitung* 612.

<sup>120</sup> S Gless and K Ligeti, 'Regulating Driving Automation in the European Union – Criminal Liability on the Road Ahead?' (2024) 15 *New Journal of European Criminal Law* 33.

<sup>121</sup> B Goodman and S Flaxman, 'EU Regulations on Algorithmic Decision Making and a Right to Explanation' (2017) 38 *AI Magazine* 50, 55 f, available at [metromemetics.net/wp-content/uploads/2016/07/1606.08813v1.pdf](https://www.metromemetics.net/wp-content/uploads/2016/07/1606.08813v1.pdf); Selbst and Powles (n 46) 235 ff.

<sup>122</sup> LM Sommerer, 'The Presumption of Innocence's Janus Head in Data-Driven Government' in E Bayamlioglu et al (eds), *Being Profiled: Cogitas Ergo Sum: 10 Years of Profiling the European Citizen* (Amsterdam University Press, 2018) 58.

## E. Privacy and Fairness Concerns – A Beneficial European Angst?

The use of AI systems in criminal matters raises manifold concerns besides the evidentiary issues discussed above, including regarding privacy and fairness, only some of which can be highlighted here. Is it right to assume that this European angst has been a beneficial driver of regulation?

In Europe, the fear of a surveillance state has prompted a multitude of domestic and European laws. As explained above (see above I.B.i), Article 8 ECHR has been seen as a bulwark against undue government infringement since the 1950s. The EU Member States first agreed on a Data Protection Directive (95/46/EC) in 1995, then proclaimed a right to protection of personal data in the Charter of Fundamental Rights of the EU in 2000 and put into effect the GDPR in 2018. The courts, in particular the CJEU, have also shaped data protection law through interpretations and rulings.

However, data processing in criminal justice has always been an exception. It is not covered by the GDPR as such, but by the LED, which addresses the protection of natural persons regarding the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences, or the execution of criminal penalties.<sup>123</sup>

The line separating civil use and criminal justice use of AI systems could nevertheless blur when, for instance, consumer product AI – such as cars – are used for fact-finding in a criminal proceeding. Both the privacy ramifications of legal acts building on expectations around privacy and the criminal justice maxim of finding the truth in criminal proceedings must be reconciled in such circumstances.

The use of facial recognition could become a paradigmatic example, as it raises a number of issues, including public–private partnerships. Facial recognition systems need huge data pools to function, which can be provided by the authorities in the form of mug shots. Creating such data pools can, however, lead to the reinforcement of bias already present in policing. Visual material could also be provided by private companies, but this raises privacy concerns if the respective individuals have not consented to be in the data pool. Data quality may also be problematic if the material lacks adequate diversity, which could affect the AI system's capability to correctly match two pictures. In the past, authorities bought pictures and services from companies that later came under scrutiny for their lack of transparency and other security flaws.<sup>124</sup> If such companies scrape photos from social media and other online sources without consent from individuals, the material cannot be used for matching, but without an adequate volume of photographs, there may be serious consequences such as wrongful identification. Similar arguments are raised regarding the use of genealogy databases for DNA sample testing by investigation authorities.<sup>125</sup> The effects of using

<sup>123</sup> LED (n 41).

<sup>124</sup> I Neroni Rezende, 'Facial Recognition in Police Hands: Assessing the "Clearview Case" from a European Perspective' (2020) 11 *New Journal of European Criminal Law* 375, 389; PI, 'Challenge Against Clearview AI in Europe', available at [privacyinternational.org/legal-action/challenge-against-clearview-ai-europe](https://privacyinternational.org/legal-action/challenge-against-clearview-ai-europe).

<sup>125</sup> S Davidowitz, '23andEveryone: Privacy Concerns with Law Enforcement's Use of Genealogy Databases to Implicate Relatives in Criminal Investigations' (2019) 85 *Brooklyn Law Review* 185.

facial recognition for criminal justice matters may be even more profound. People might feel safer overall if criminals are identified, but also less inclined to exercise legal rights that put them under the gaze of the authorities, such as taking part in demonstrations.<sup>126</sup>

### III. Résumé and a Look into the Future

The digital era changes criminal justice, just as it has reshaped other areas of the law. However, the exact nature of the impact AI systems will have on criminal trials is still unclear. Focusing on evidentiary issues, Europe could again take a leading role in regulation as it has in the field of data protection as well as fundamental rights with its two institutional stakeholders, the CoE and EU.

For guidance on how to regulate cybernated fact-finding, it can draw on an emerging body of law, including on the one hand ECtHR case law and the Budapest Convention, as well as other CoE instruments, and, on the other, the relevant EU acts, including the GDPR and the EU AI Act. These different legal acts have the potential to intertwine and establish a grid for building an adequate regulation of AI-generated evidence. As things stand, however, the body of European legislation regulating AI requires further adaptation to rise to the challenges of AI use in criminal justice in the future. Particularly, the AI Act's potential to regulate risks can be used to address challenges in the criminal justice systems. The mix of product safety and fundamental rights protections holds the promise for a standardised approach of regulating AI employment in forensics.

When regulating AI-generated evidence it is important to reflect the classic balance of making all relevant information accessible for fact-finding in criminal proceedings and, at the same time, adequately addressing the novel risks of such evidence regarding reliability and fairness of fact-finding in judicial procedures. Including AI-generated evidence in criminal proceedings will raise specific issues, such as balancing mechanisms for conflicts arising from the right to privacy in a criminal investigation that integrate Consumer Product AI and everyday items that monitor their users' behaviour; these issues cannot be resolved by existing legislation.<sup>127</sup> For instance, the possible inclusion of AI-generated evidence from Consumer Product AI as evidence is not clearly regulated by the EU AI Act, which may only apply to explicitly forensic tools; neither does the LED offer clear protection for data processed during criminal proceedings. Coupled with the issue of fairness raised by the omission from the EU e-evidence package of standards of proof for AI-generated evidence and as the general opacity of AI systems, defendants and the general public

<sup>126</sup> K Hamann and R Smith, 'Facial Recognition Technology: Where Will It Take Us?' (Prosecutors' Center for Excellence, 20 May 2019) 11–13, available at [pceinc.org/facial-recognition-technology-where-will-it-take-us](http://pceinc.org/facial-recognition-technology-where-will-it-take-us); JW Penney, 'Understanding Chilling Effects' (2022) 106 *Minnesota Law Review* 1452.

<sup>127</sup> S Gless, 'AI in the Courtroom: A Comparative Analysis of Machine Evidence in Criminal Trials' (2020) 51 *Georgetown Journal of International Law* 195, 209 and 216.

are left in a difficult position unless there are improvements to existing legislation to target these issues.

However, despite the potential dangers posed by admitting AI-generated evidence into criminal proceedings, the promise for a comprehensive fact-finding is clear. Embracing and properly regulating AI could lead to a future with a potentially huge body of evidence that is easily transferrable across borders. With a proper taxonomy of AI-generated evidence adopted across Europe, many of the issues surrounding reliability could be mitigated. And, with improvements to data privacy laws to account for the unique methods of AI systems, the public–private divide could be maintained without posing risks to the general public through over-reaching predictive policing or FRT. AI-generated evidence will continue to grow in importance and Europe, with the CoE and EU as its primary legislative stakeholders, has laid the first foundations to begin to control it. It must now resolutely continue on this path.