Himachal Pradesh National Law University, Shimla (India)

A UGC CARE Indexed Law Journal

Journal Articles | *ISSN:2582-1903* | *Shimla Law Review*

Volume: VII (2024)

AI GENERATED EVIDENCE BROUGHT BEFORE EUROPEAN COURTS: Further Activation of Evidence Law, Defence Rights, and Human Oversight Needed

Paul De Hert & Georgios Bouchagiar

This article can be downloaded from:
https://www.hpnlu.ac.in/journal-article.aspx?ref-id=47&article-id=628

Recommended Citation:
Paul De Hert & Georgios Bouchagiar, *AI Generated Evidence Brought Before European Courts: Further Activation of Evidence Law, Defence Rights, and Human Oversight Needed* VII SML. L. REV. 1 (2024).
DoI: https://doi.org/10.70556/hpnlu-slr-v7-I1-2024-01

This Article is published and brought to you for free and open access by Himachal Pradesh National Law University, Shimla. For more information, please contact *editorslr@hpnlu.ac.in*

Contents

Volume VII | *ISSN: 2582-1903* | April 2024 - March 2025

Special Article *Page*

1. AI GENERATED EVIDENCE BROUGHT BEFORE EUROPEAN COURTS: Further Activation of Evidence Law, Defence Rights, and Human Oversight Needed
   Paul De Hert & Georgios Bouchagiar 1

Articles

2. 'ONE NATION ONE ELECTION': A Juridical Critique of the Constitution (One Hundred and Twenty-Ninth Amendment) Bill, 2024 41
   Virendra Kumar
3. ELECTORAL HARMONY AND CONSTITUTIONAL BALANCE: A Constitutional and Theoretical Analyses of India's Quest for One Nation One Election 65
   Tamesh Kumar Pandey & Priti Saxena
4. THE ELECTION COMMISSIONERS' APPOINTMENT ACT, 2023: Fulfilment of a Constitutional Expectation or a Device to Maintain the Status Quo? 92
   Monica Chaudhary
5. LAW AS THE ONTOLOGY OF 'ONE' AND 'CONSTITUTIONAL REASON': India's 'One Nation One Election' Plan, Democracy, Federalism, and Electoral Justice 128
   Chanchal Kumar Singh & Mritunjay Kumar
6. EFFICACY OF SEXUAL HARASSMENT OF WOMEN AT WORKPLACE (PREVENTION, PROHIBITION, AND REDRESSAL) ACT, 2013: A Critical Analysis 213
   Aayush Raj
7. POSTHUMANISM: Challenges for Contemporary Legal and Social Theories 244
   Chanchal Sharma
8. BEYOND THE UNIFORM CIVIL CODE DEBATE: Critiquing Goa's Succession Law 262
   Shukr Usgaokar

9. INDIA'S DECRIMINALISATION POLICY: The Past, Present, and Future of Governance, Risk Management, and Compliance under the Companies Act, 2013
Harpreet Kaur 293

Notes and Comments

10. ARTIFICIAL INTELLIGENCE INVENTIONS AND THEIR PATENTABILITY
Anita Yadav & Rabikant Pandey 312

11. 75 YEARS OF THE PRACTICE OF EXTRADITION LAWS IN INDIA: Evolution, Expectations, and Realities
Namita Vashishtha & Shourie Anand Singh 338

12. *IN RE:* SECTION 6A OF THE CITIZENSHIP AMENDMENT ACT, 1955: A Comment
Muskan & Prachi 370

13. RIGHTS AND ENTITLEMENTS: A Commentary on Chapter II of the RPwD Act, 2016
Ashpinder kaur & Rimjim Kheda 380

14. FISCAL FEDERALISM FROM PYRAMID TO MATRIX: An Analysis of the Supreme Court Judgment in *Mineral Area Development Authority* v. *Steel Authority of India* (2024 S.C.C. OnLine S.C. 1796)
Arya Kumari 394

AI GENERATED EVIDENCE BROUGHT BEFORE EUROPEAN COURTS: Further Activation of Evidence Law, Defence Rights, and Human Oversight Needed

Paul De Hert[*] *& Georgios Bouchagiar*[**]

[Abstract: *Artificial Intelligence (AI) is used in the criminal justice system to support human decision-making at various stages of the proceedings, despite criticism on its opacity, complexity, non-contestability or ability to trigger unfair discrimination. This contribution assesses such AI-implementations under the light of European evidence law, as well as the right to a fair trial, non-discrimination and the protection of personal data. Europe only has minimum rules in the criminal law sphere, making a more generic assessment difficult. Our case law and literature analyses show promising signs of greater awareness of certain risks created by this growing use of AI in the criminal justice chain; there is a clear need for strengthening reliability requirements within evidence law; strengthening human oversight; and strengthening defence rights, including the right to contradict and control expert outcomes.*]

I

Introduction: AI, Defence Rights, Fundamental Rights and Legality[1]

We are entering an era, where proprietary and black-box AI systems, generating or processing evidential materials, are used in criminal courtrooms to support judicial decision-making.[2] AI analyses large amounts of data in a fast and sophisticated way

[*] Prof. Paul De Hert's work addresses problems in the area of privacy/technology, human rights and criminal law. He is Full Professor at the Vrije Universiteit Brussels (VUB-LSTS) and Associated-Professor at Tilburg University (TILT). Email: paul.de.hert@tilburguniversity.edu

[**] *Dr. Georgios Bouchagiar is a postdoctoral researcher on law and technology at the Free University of Brussels (LSTS). He holds a PhD on surveillance and privacy, an LLM on law and technology, a MSc on intellectual property and new technologies and a law degree. Email: georgios.bouchagiar@vub.be*

[1] The authors would like to thank Francesco Paolo Levantino and Andreas Kanakakis for their constructive feedback and comments on initial drafts of this paper.

[2] *See*, *State* v. *Puloka* (No 21-1-04851-2KNT), where the AI-enhanced video, adding its own content *via* opaque methods, was excluded. *See also*, *State* v. *Arteaga* (Superior Court of New Jersey, Appellate Division; Docket No A-3078-21, Decided on 7 June 2023), favouring the defence regarding discovery for an opaque (possibly untested and biased) face recognition tool. For a

Contd…

to deliver outputs (like a decision, a likelihood or content). It is said to be particularly attractive to law enforcement and criminal justice authorities that can be assisted in investigating crime, charging or convicting suspects or evaluating defendants at various stages of the criminal proceedings (*e.g.*, at the pretrial, the sentencing or the post-sentencing phase). Despite critical remarks on AI's lack of transparency and intelligibility, its risk of unfairly discriminative impact and its error rates, due to, among others, bias or inadequate review, training and validation, the use of AI has been embraced by law enforcement and criminal justice authorities tempted by the technology's promises.

An extensive analysis of the impact of AI on European procedures has been done by Francesca Palmiotto in her 2023 PhD 'Artificial Intelligence and the Transformation of Criminal Trials: Preserving Fairness in Europe'.[3] Palmiotto seems to perceive AI as a game changer, in particular for inquisitorial systems. She contends that the use of AI systems to generate evidence in criminal trials leads to a threefold transformation: one, new actors enter the scene, bringing commercial interests into the proceedings; two, criminal trials become increasingly mediated by experts and; three, the traditional model of evidence examination is fundamentally changed, that is, from inside the courtroom, *audi alteram partem*, to outside the trial, in *absentia* of the parties.

Our contribution assesses whether and the degree to which AI-implementations can respect fundamental principles that must be in place at various stages of the criminal procedure. It adds an update literature analysis to Palmiotto's work and takes into account the novel developments in the United States' (US) legal system. After a glance at these US developments (section II), some general observations on European criminal procedure (adversarial/inquisitorial) are made (section III). The paper then looks at several European legal frameworks on evidence (sections IV-VII); the right to a fair trial (section VIII); non-discrimination (section IX); and the legality principle (sections X-XI).

We find only basic guidance for correct use of AI-evidence in the European case law of the European Court of Human Rights (ECtHR) and the Court of Justice of the European Union (CJEU) characterised by self-restraint and sensitivity to interfere with the organisation of criminal procedure in a Continent characterized by a

positive, but cautious European analysis, *see*: Helga Molbæk-Steensig & Alexandre Quemy, *Artificial Intelligence and Fair Trial Rights*' ARTIFICIAL INTELLIGENCE AND HUMAN RIGHTS ,Alberto Quintavalla & Jeroen Temperman, 265-280 (2023).

[3] F. Palmiotto, ARTIFICIAL INTELLIGENCE AND THE TRANSFORMATION OF CRIMINAL TRIALS: PRESERVING FAIRNESS IN EUROPE (PhD Thesis, European University Institute, 2023), *available at*: https://cadmus.eui.eu/handle/1814/75243 (last visited 24 Oct., 2024). On the impact of AI in the mid-term future, *see also*, B Van Der Sloot, *De impact van AI op de rechtspraak: Is het bewijsrecht klaar voor de 21ste eeuw?* RECHTSTREEKS 33-40 (2024), *available at*: https://repository.tilburguniversity.edu/server/api/core/bitstreams/c5451dc3-04bc-4da1-8e77-ae96e65f266b/content (last visited 12 June, 2025).

mixture of legal regimes (inquisitorial, adversarial, or mixed). The overall picture in Europe remains one of fragmentation and pluralism, partly because the ECtHR is not always very activist and partly due to the presence of both inquisitorial and adversarial systems in Europe. Throughout our analysis we highlight diverse proposals in European literature to recognize more explicit novel rights in evidence law and criminal procedure law on fair trial.

The 2016 European law on data protection for law enforcement authorities (LED) and the 2024 EU AI Act add unimportant elements with regard to principles and enforcement. The example is of course the LED with new important processing principles (*e.g.,* data quality, minimisation, human oversight, purpose limitation or notification rights) and a prime role for a new actor in the criminal law sphere: the data protection authorities. But it is hard to believe that both laws will make a difference in court, due to their lack of concreteness.

Recognition of novel defence rights and codification of novel interpretations of existing fair trial rights can and must address transformations of the traditional procedure caused by AI and preserve the core principles of a fair trial. These clarifications are of utmost importance, given the need to ensure impartiality and autonomy of judges, fairness of trials, as well as to uphold human and procedural rights, such as, in the adversarial principle, the principle of equality of arms, effective remedy and reasoned ruling. In our discussion, we include the United Nations' (UN's) guidance (toolbox) on AI and a CJEU case on AI (*Ligue des Droits Humains* v. *Conseil des Ministers*)[4] (sections XII-XIII). We close the discussion by comparing the European framework(s) with the US legal framework (section 14) and with recommendations.[5]

[4] *Ligue des Droits Humains* v. *Council of Ministers*, ECLI:EU:C:2022:491.

[5] Before proceeding to the analysis, it is worth making a clarification on and distinction between the Europe of the Council of Europe (CoE) and the Europe of the European Union (EU). The former includes a larger number of addressees, the Contracting Parties; and its main legal instrument on human rights protection is the 1950 European Convention on Human Rights (ECHR), interpreted by the European Court of Human Rights (ECtHR). The EU has a smaller audience, the Member States; its key tool on human rights protection is the 2000 Charter of Fundamental Rights (CFR) that came into force in 2009; and the Court of Justice of the European Union (CJEU) supervises compliance with the EU law. The analysis of this chapter mainly refers to the Council of Europe (except where otherwise indicated); but it also provides for adequate footnoting on corresponding provisions of the EU legal framework, which has been inspired by and has heavily relied upon the CoE's framework. On the European legal framework, *see*, in general: P Craig & G de Búrca, EU LAW: TEXT, CASES, AND MATERIALS, 6th edn, Oxford University Press (2015); R. Schütze, AN INTRODUCTION TO EUROPEAN LAW, 2nd edn, Cambridge University Press (2015).

II

AI-Generated Evidence in Court: Short Summary of Recent US Developments

Probably, no legal system in the world is more complete on evidence law than the American system of criminal procedure. The constitutional basis and the regulatory framework on criminal procedure and evidence law, before the AI era, had already been impressive in their strive for detail and the need to bring evidence in the court that can be controlled and understood (by a jury!). A previous paper highlighted the benefits of the American adversarial system for challenging AI-uses in court; and it underlined its vigilance: the Federal Rules on Evidence (as amended in 2023) and other regulatory developments together with advances in case law towards better judicial gatekeeping with regards to AI evidence in court; and this, in the spirit of the legacy of the US criminal system, seems to have put an end to an initial period of obliging acceptance of AI evidence in court, despite strict rules on evidence.[6] Every use of AI in the criminal justice system is expected to be compliant with evidence-related law, as well as key constitutional rights, in particular, the right to due process, equal protection and privacy. The American legal landscape seems to fulfil this requirement, at least on paper. Our analysis showed that legal practice is far from perfect and that challenges remain,[7] but we agree with European

[6] P De Hert & G Bouchagiar, *Specific Laws Governing Use of AI in Criminal Procedure and Attentive Criminal Judges: American Songbook for Global Listeners*, 6 SLR 1-31(2024); P De Hert & G Bouchagiar, *Law Enforcement Relevant Regulation, Attentive Criminal Courts and AI. American Songbook for European Listeners,* 10 BRUSSELS PRIVACY HUB 9 (2024), *available at*: https://brusselsprivacyhub.com/wp-content/uploads/2024/12/Law-Enforcement-Relevant-Regulation-Attentive-Criminal-Courts-and-AI.-American-Songbook-for-European-Listeners-v2.pdf (last visited 5 Dec., 2024)., where we discussed cases, such as *State* v. *Puloka* (No 21-1-04851-2KNT) and *State* v. *Arteaga* (Docket No A-3078-21), as well as regulatory interventions, like the Federal Rule of Evidence 702 (as amended in 2023), AI Policy 2024 [Executive Office of the President (Office of Management and Budget), 'Memorandum for the Heads of Executive Departments and Agencies' (Executive Office of the President, 28 March 2024) *available at :* https://www.whitehouse.gov/wp-content/uploads/2024/03/M-24-10-Advancing-Governance-Innovation-and-Risk-Management-for-Agency-Use-of-Artificial-Intelligence.pdf (last visited 26 Nov., 2024) and California's Rules of Court, 2024. *See also*, discussions on revising the Federal Rule of Evidence 901 (governing the authentication and identification of evidence in legal proceedings), with a view to overcoming concerns posed by AI-generated evidence and deepfakes: J Bickford, *AI Is Coming, But the Rules Aren't Ready*, GEORGETOWN LAW TECHNOLOGY REVIEW (2025), *available at :* https://georgetownlawtechreview.org/ai-is-coming-but-the-rules-arent-ready/GLTR-01-2025/ (last visited 12 June 2025); P Langham, *Amendments to the Federal Rules of Evidence Under Consideration to Address Deepfakes and Voice Cloning*, MARYLAND STATE BAR ASSOCIATION, 22 May 2024), *available at :* https://www.msba.org/site/site/content/News-and-Publications/News/General-News/Amendments_to_the_Federal_Rules_of_Evidence_Under_Consideration_To_Address_AI_and_Deepfakes.aspx (last visited 12 June 2025).

[7] We quoted authors, arguing that testing expert evidence in court in the context of AI is still not

Contd…

comparative scholar, Sabine Gless, who is positive about the readiness for AI of the US criminal procedure.[8] More than Europe, the US legal system (including its case law) has explored the question of what constitutes scientific knowledge and when an expert in court can be trusted. For Gless, European inquisitorial systems tend to blur what she calls 'the science dilemma'. Europeans know what science is when they see it.[9]

American literature, at least, is very aware of the pitfalls of uncritical reliance on science and technology and argues that AI-uses that are not comprehensible by humans should be avoided, while making critical decisions. This follows from evidence-related standards that demand the involvement of an adequately qualified human expert, who can assist the court in understanding evidence or determining a matter. It equally follows from the due process requirement to give reasons for decisions that should, moreover, be contestable.

For Duke professor, Brandon Garrett, the need for black box AI systems in the legal system is overstated. Interpretable AI systems perform as well and have clear advantages in terms of due process rights.[10] His presumption against black box AI is not absolute, since there can be more openness towards opaque AI-implementations in less critical contexts, where statistical precision is required and where no defence rights or other human rights are at stake (*e.g.*, uses by law enforcement to merely improve resolution of a low-quality video-footage). In such

---

perfect, without consistent forensic standards, external validation and industry standards. We also discussed authors, who question the readily acceptance of black box AI by law enforcement and call for judicial rulings and legislation to safeguard a right to interpretable forensic AI and stop using black box AI.

8 S Gless, *Rules on Expert Testimony based on a Comparative Perspective on Device Evidence*, The Advent of AI: Reshaping Criminal Procedure, University of Luxembourg, (7-8 November 2024) *available at:* https://aiandcriminaljustice.uni.lu/2024/11/06/presentations-conference-crim_ai/ (last visited 12 Nov., 2024). *See* also: S Gless, F Lederer & T. Weigend, *AI-Based Evidence in Criminal Trials?* , 59(1) TULSA LAW REVIEW 1-36(2024); S Gless, *AI in the Courtroom: A Comparative Analysis of Machine Evidence in Criminal Trials*, 51 GEORGETOWN JOURNAL OF INTERNATIONAL LAW 195-253 (2020).

9 The American approach is more thoughtful and elaborated. On the negative side, adversarial systems, like in the US regime, Gless observes, disregard the cost problem: one needs resources to bring in science and experts; resources that many Americans do not have.

10 B Garrett, *Rule 702 Amendments and their Impact on Admissibility of AI Evidence The Advent of AI: Reshaping Criminal Procedure*, University of Luxembourg, (Nov., 7-8 2024) *available at:* https://aiandcriminaljustice.uni.lu/2024/11/06/presentations-conference-crim_ai/ (last visited 12 Nov., 2024); B Garrett & C Rudin, *Interpretable algorithmic forensics,* 120(41) PROC NATL ACAD SCI USA (2023); B L Garrett & C Rudin, *The Right to a Glass Box: Rethinking the Use of Artificial Intelligence in Criminal Justice,* 109 CORNELL LAW REVIEW 561 (2024). The burden rests on the government to justify any departure by law enforcement (or judges) from the norm that all lawyers, judges, and jurors can fully understand AI. The presumption should be in favour of glass box AI (designed to be interpretable), absent strong evidence to the contrary. Together with his co-author, he calls for national and local case law and regulation to safeguard, in all criminal cases, the right to glass box AI.

less critical cases, however, the goal pursued by a given AI-use should be limited to corroborating a human-made decision that should, moreover, be independently supported by other materials;[11] and AI's overall performance should be subjected to enhanced checks and balances, ensuring that transparency and other crucial safeguards are in place.[12]

Contrary, Van Banning argues that, despite limited comprehensibility of AI-implementations, there can be control, with more and more AI-uses in the criminal justice arena, and new safeguards and guarantees, like validation to assess an AI-system's proper functioning or certification mechanisms, could be the way forward for law enforcement.[13] Something that does not negate (rather seems to support) Garrett's position on glass box AI.

Most American authors agree that higher standards on reliability and expert evidence are needed. More standard-setting measures could ensure higher levelled-reliability by solely permitting tools that have been adequately trained on sufficient and accurate data, as well as tested, peer-reviewed and generally accepted by the relevant scientific community (as required by expert evidence-related provisions and the *Daubert*-standard).[14] Both Sabine Gless and Jumana Musa are critical of

---

[11] On judicial discretion concerning the use of technologies for corroboration goals, see: *Loomis*. Of relevance to judicial discretion in considering AI are automation bias that could lead to over-reliance upon the technology, probably without judges being aware of their over-trust on the technology. On automation bias, *see*: K Freeman, *Algorithmic Injustice: How the Wisconsin Supreme Court Failed to Protect Due Process Rights in State v. Loomis*, 18 NORTH CAROLINA JOURNAL OF LAW & TECHNOLOGY 75, 98 (2016).

[12] In this regard, *see*: D Barysė & R Sarel, *Algorithms in the Court: Does It Matter Which Part of the Judicial Decision-Making is Automated?* 32 ARTIFICIAL INTELLIGENCE AND LAW 117 (2024), discussing trust in judges' ability to analyse information and confidence in algorithmic implementations for information-gathering purposes. Transparency is imperative for the defence to effectively contest allegations against it and challenge their reliability, under the general requirements of due process. It is also of major importance when it comes to data processing in the fact-finding or other context. To increase transparency of AI-uses, it would be preferable to: use open or free tools, whose modus operandi can be accessible and reviewable; maximize accessibility, in case of IP-protected AI, via procedural legal instruments (like protective orders) granting conditional access without disproportionately limiting Intellectual Property (IP) rights; and offer access to ancillary materials (such as underlying methodologies or mathematics), whose consideration could assist experts in comprehending the modus operandi of a given processing operation.

[13] J Van Banning, THE AI ACT AS A SAFEGUARD FOR EQUALITY OF ARMS IN DATA-DRIVEN CRIMINAL INVESTIGATIONS, 201, 202 (2025), ('(…) *The specific character of AI tools used for handling evidence, it is argued, do not allow for easy contestation during trial, possibly making other mechanisms for safeguarding fair trial rights more appropriate, like the use of certification mechanisms* (…)'), 205, 207 ('(…) *Even if one does not completely understand the way the tool functions, there are still a number of ways to attempt to assess its use. For example, any system used in forensics, even if it is a black box, can be validated in order to assess whether it functions properly* (…)').

[14] The use of such tools could eliminate or reduce bias and error rates, enhance probative value of AI and make it more relevant. Such standard-setting could include permissible factors and inputs that could promote accuracy; and exclude non-permissible factors/inputs could result in

Contd…

testing expert evidence in US courts in the context of AI for the reasons that there is no consistent forensic standards, a lack of external validation, as well as industry standards.[15]

Another battle in the US discussions is on the level of scrutiny in cases of possible discrimination due to the use of AI. Critical voices hold that when the defence demonstrates to the court that it is most likely that AI has caused discrimination, strict scrutiny should be applied. Given that, in the criminal justice system, a wide array of fundamental rights may be involved, but also interfered with by opaque AI-implementations, it would be desirable to apply the strict scrutiny standard, when determining whether a given AI-use can lead to unfair discrimination. Such scrutiny could require a heavy burden of justification and necessity of the relevant AI-use to achieve the goal pursued. Furthermore, given that AI could autonomously go beyond and escape from the intentions of its developer, it would be fair to solely focus on 'discriminatory impact' and *not* require proof of discriminatory intent.

It is a matter of the legality principle to ensure that people are aware of and can foresee what is permitted and what is prohibited, how governments operate and what kind of governmental actions limit their rights and freedoms. *Above*, we discussed Brandon Garrett's proposal for regulation to safeguard in all criminal cases the right to glass box AI. The message that more laws are needed, also in the US, is picked up. New regulatory developments in the US (both at federal and state level), the Federal Rule of Evidence 702 (as amended in 2023), the 2024 AI Policy and California's 2024 Rules of Court; impose concrete obligations with a view to protecting those who might become vulnerable to AI-uses.[16] They demonstrate that

---

unfair discrimination. In this regard, see the recommended reliability validation enabling framework for the evaluation of digital forensics in criminal investigations: R Stoykova & K Franke, *Reliability Validation Enabling Framework (RVEF) for Digital Forensics in Criminal Investigations*, 45 FORENSIC SCIENCE INTERNATIONAL: DIGITAL INVESTIGATION 301554 (2023), *available at:* https://doi.org/10.1016/j.fsidi.2023.301554 (last visited 8 Sept., 2024). Such standard-setting, aimed at minimising undue discrimination, appears imperative, especially in light of current un/under-regulated AI uses, possibly resulting in bias, when making critical decisions. The reality of bias in this context is without dispute. In this regard, *see*: Robert Booth, *Revealed: bias found in AI system used to detect UK benefits fraud*, THE GUARDIAN (Dec. 6, 2024), *available at:* https://www.theguardian.com/society/2024/dec/06/revealed-bias-found-in-ai-system-used-to-detect-uk-benefits (last visited 9 Dec., 2024).

[15] S Gless, *Rules on Expert Testimony based on a Comparative Perspective*, *available at:* https://aiandcriminaljustice.uni.lu/wp-content/uploads/sites/260/2024/11/Gless_presentation_Luxembourg-Nov-2024_final.pdf (last visited 2 Nov., 2025); J Musa, *Challenges Raised by AI Evidence: Lessons from the US Criminal Justice System, The Advent of AI: Reshaping Criminal Procedure*, University of Luxembourg (Nov., 7-8 2024) *available at:* https://aiandcriminaljustice.uni.lu/2024/11/06/presentations-conference-crim_ai/ (last visited 12 Nov., 2024).

[16] This probably includes a rather large number of people, given the possible intrusiveness of contemporary AI implementations. On vulnerability, *see* G Malgieri, VULNERABILITY AND DATA PROTECTION LAW, Oxford University Press (2023).

detailed and clear lawmaking can promote the people's general interest in knowing and foreseeing state actions and their limits. The 2024 AI Policy sets out a rigorous baseline for responsible AI-implementations and imposes concrete obligations on various authorities, including law enforcement, ranging from transparency obligations (for instance, on risk assessments)[17] to independent scrutiny (such as impact assessments concerning policing).[18]

Our paper now turns to Europe with its characteristic diversity of criminal law regimes (countries like the United Kingdom with an adversary system; continental states with a dominantly inquisitorial system; mixed systems, like in Italy). We look at evidence law, due process rights, protection against discrimination and regulatory frameworks of all sorts that may have an impact on the use of AI.

III

Criminal Procedure in Europe: General Oversight

It can be generally stated that every type of criminal procedure is based on either an accusatory (common law) or an inquisitorial (civil law) procedural model.[19] The US is an example of the former; and most European continental systems, like France, are examples of the latter. The difference between the two models lies in the horizontal and vertical structure of the administration of criminal justice. In the accusatory law suit, two equal and autonomous parties (the suspect and the prosecuting parties) make their own case concerning the alleged criminal offences before a neutral judge; while, in the inquisitory model, an official authority collects evidence, on its own initiative, without consulting either party, and uses this evidence to bring the truth to light. In brief, the following general attributes of the accusatory criminal law process can be pointed out:[20]

---

17 Executive Office of the President (Office of Management and Budget), 'Memorandum for the Heads of Executive Departments and Agencies' (n 5), Appendix I, point 2.

18 Executive Office of the President (Office of Management and Budget), 'Memorandum for the Heads of Executive Departments & Agencies' (n 5), section 5 (Managing risks from the use of artificial intelligence).

19 T Decaigny & P De Hert, *You Can Change the Colour of Your Hair, Not Your Hair. Converging Is What Inquisitorial and Adversarial Systems Rarely Do*, 75 BOOM LEMMA UITGEVERS 235 (2013); T Decaigny, *Inquisitorial and Adversarial Expert Examinations in the Case Law of the European Court of Human Rights*, 5(2*)* NEW JOURNALS OF EUROPEAN CRIMINAL LAW 149-166 (2014).

20 Helga Molbæk-Steensig & Alexandre Quemy, ARTIFICIAL INTELLIGENCE AND FAIR TRIAL RIGHTS, 265-280(2023); B de Smet, DE HERVORMING VAN HET STRAFRECHTELIJK VOORONDERZOEK IN BELGIË, 12-23(1996). *See* also: B de Smet, INTERNATIONALE SAMENWERKING IN STRAFZAKEN TUSSEN ANGELSAKSISCHE EN CONTINENTALE LANDEN, 19-43 (1999); F Tulkens, CRIMINAL PROCEDURE: MAIN COMPARABLE FEATURES OF THE NATIONAL SYSTEMS, 8-9(1995); T Thaman, *Trial by Jury and the Constitutional Rights of the Accused in Russia*, 4 (1) EAST EUROPEAN CONSTITUTIONAL REVIEW 77-80

Contd…

Firstly, within the accusatory tradition, the belief does not exist that the objective truth lies in evidential materials. Both parties must justify their truth (through evidential materials); and the judge or the jury decide which of the stories brought forward is the most believable and best mirrors the truth.

Secondly, the accusatory procedure leans on a liberal theory of basic rights (procedural guarantees are no goal in themselves, they only extend to protect the suspect from arbitrary government actions), which gives the parties an important side effect, *i.e.*, the chance to come to an agreement themselves, to abandon the lawsuit and to agree upon a certain result through negotiations.[21]

Thirdly, the furnishing of proof is usually based on the so-called immediacy principle, which implies that the judge or the jury only use the evidence laid before them by the parties to support their inner convictions and do not use, or use to a lesser extent, the evidence collected by the police in the pre-trial phase which could be biased (hearsay).

Fourthly, the collection of evidential materials is the complete responsibility of the parties themselves; and the accused, who is considered a full party and can decide him/herself whether he/she wants to actively contribute to 'his'/'her' truth (process autonomy).

Lastly, coercive measures can only be employed with the authority of the judge.

Contrary to this accusatory procedure, the inquisitory procedure draws from the idea of an objective, material truth that 'must' be found through a procedure where a professional judge forms the pivotal figure, where regulations and procedural stipulations are binding, and the pre-trial investigation carries a determined weight. The Belgian, French and Dutch systems are generally described as having systems of moderate inquisitory judicial procedures. This system distinguishes itself by a strict division between the pre-trial and the trial investigation, where the exercise of collecting evidence is done by the public prosecution's department or an independent magistrate (the examining magistrate). The department of public prosecutions is involved in both the investigation and the prosecution.[22]

(1995). For a more legal theoretical account of the difference in legal culture, *see*: B Edelman, UNIVERSALITY AND HUMAN RIGHTS, 97-107(1995); F Macchiarola, *Finding the Truth in an American Criminal Trial: Some Observations*, 85-87(1997).

[21] Usually, the result of these negotiations is that the suspect foregoes the trial in exchange for a promise from the prosecuting party to demand a mild punishment. On the success of the guilty plea and plea bargaining in English and American trial systems: B de Smet, DE HERVORMING VAN HET STRAFRECHTELIJK VOORONDERZOEK IN BELGIË, 12-23(1996).

[22] Another institution is the '*juge d'instruction*'. This examining magistrate is expected to conduct an official investigation into the alleged punishable crimes that should contain both incriminating and exculpatory evidence. The examining magistrate must also investigate the legality and opportunism used in the methods of coercion. These can be ordered by him/her on his/her own initiative, that is to say without a formal request from the prosecuting party.

At trial level, Europe lacks uniformity in various respects (from evidence law (see *below*) to jury trials). There is a right to a jury trial in most countries, even those belonging to the inquisitorial tradition; albeit there are countries that fully rely upon professional magistrates, like the Netherlands. The difference with the US is obvious: in the US, the right to *trial by jury* is laid down in the constitutions of the states and the federal constitution. But the countries without any trace of lay participation in Europe are only few.[23] In England, the presence and the role of the jury is fundamental in criminal procedure.[24] This aspect of the organisation of the trial is not a question of historical anecdotes. The presence of lay judges has consequences for the organisation of the trial and the evidence. They do not have a criminal file prepared by the police and the prosecutor. There is simply no file at the beginning of the trial. All evidence must be brought to court and be understandable to the jury (in terms of use of AI, this is an enormous advantage for the defence).

These aspects of comparative criminal procedure have concrete impact on the reception of AI in court. Evidence that is considered unreliable or illegal will not make it into court but will be excluded in the American system based on the jury system. In the inquisitorial systems that work with bench judges, *all* evidence will be seen by them, who will then have to justify their use of it (if they use it).[25]

IV

AI-Generated Evidence and Evidence Rules on Reliability in Europe

For evidence to be acceptable or admissible, it firstly needs to be *reliable*, meaning meeting the requirements of authenticity and integrity.[26] Reliability of AI-generated

---

[23] The basic right to a jury is also well established in several continental law systems: J Pradel, *Procédure pénale comparée dans les systèmes modernes: Rapports de synthèse des colloques de l'ISISC*, 15 NOUVELLES ETUDES PÉNALES 44 (1998), and is, besides, experiencing a certain constitutional revival. Cf S Thaman, *The Jury Returns to Continental Europe: Russia and Spain Return to the Classic Jury as a Catalyst in a Move to a more Adversary Form of Criminal Trial*,17 NOUVELLES ETUDES PÉNALES 343-363(1998).

[24] There is always the possibility for the defence to make an appeal to the Crown Court that resides with a jury for judgements passed on smaller criminal offences that are brought before the Magistrates' Court which resides without a jury. *See*: A Verbeke, ELEMENTEN VAN ANGLO-AMERIKAANS RECHT. EEN SELECTIE VAN ENGELSE EN AMERIKAANSE JURIDISCHE TEKSTEN GESITUEERD IN HET BETROKKEN RECHTSSYSTEEM 109-111, (1994).

[25] *Supra* note 15.

[26] The ECtHR has established certain requirements for reliability: there must be no doubt on authenticity and integrity of evidence. Authenticity refers to the irrefutable provenance of evidence. Integrity refers to the fact that evidence remains intact and that it is not tampered with during its collection or subsequent handling. Moreover, reliability is inextricably linked to the weight of evidence in a criminal trial. *See*, K Quezada-Tavarez, P Vogiatzoglou & S Royer, *Legal*

Contd…

evidence can be affected at several levels: the authenticity and integrity of the 'raw' digital data can be at stake (think about data given by informants); the AI tools can contain miscodes that could remain hidden because of black box-opacity; and human experts operating AI tools can jeopardise reliability of AI evidence.

US law is very explicit on reliability and judges are given considerable discretion to check these requirements, acting as gatekeepers for adequately reliable evidence.[27] In her comparative study, Sabine Gless finds that many European codes on criminal procedure, particularly those of continental legal systems, lack similar explicit requirements.[28] In practice, the concern for reliability is more than present. Most European states share the same basic, fundamental notions of evidence law with only slight differences in terminology or phrasing.[29] Concepts, like 'reliability' (connected with authenticity and integrity) and related notions, such as 'relevance', 'credibility' and 'validity',[30] are

Challenges in Bringing AI Evidence to the Criminal Courtroom, 12(4) NEW JOURNAL OF EUROPEAN CRIMINAL LAW 531, 541, (2021), N Sunde & I E Dror, *A hierarchy of expert performance (HEP) applied to digital forensics: Reliability and biasability in digital forensics decision making*, 37 FORENSIC SCIENCE INTERNATIONAL: DIGITAL INVESTIGATION(2021).

27 According to the US Federal Rules of Evidence,2024 information is admissible, if it is 'relevant'– having a tendency to prove a fact. Relevant materials can be excluded, if their probative value can be substantially outweighed by, among others, unfair prejudice or misleading. *See* Federal Rules of Evidence 402 and 403, Federal Rule of Evidence 401; Notes of Advisory Committee on Federal Rule of Evidence 401, citing California Evidence Code S. 210 that in turn defines 'relevant evidence' in terms of its tendency to prove a fact.

28 *Supra* note 1; S Gless, F Lederer & T Weigend, *AI-Based Evidence in Criminal Trials?*, 59 TULSA LAW REVIEW 1-37 (2024); S Gless, *AI in the Courtroom*, 51 GEORGETOWN JOURNAL OF INTERNATIONAL LAW 195-253(2020).

29 Elsewhere, we have addressed in more detail evidence-related matters in particular relation to AI. *See:* P De Hert, A Kanakakis & J Sajfert, '*Evidence Law, Data Protection Standards and the Use of AI for Investigation and Prosecution*' in K Ligeti & N Visco (eds), AI EVIDENCE AND CRIMINAL PROCEEDINGS (Hart Publishing forthcoming).

30 R Glover, MURPHY ON EVIDENCE,39-42(2017); D Nicolson & D Auchie, *Assessing Witness Credibility and reliability: Engaging Experts and Disengaging Gage?* in SCOTTISH CRIMINAL EVIDENCE LAW: CURRENT DEVELOPMENTS AND FUTURE TRENDS 161-193 (P Duff & PR Ferguson, eds. 2017); K Quezada-Tavarez, P Vogiatzoglou & S Royer, *Legal Challenges in Bringing AI Evidence to the Criminal Courtroom*, 12(4) JOURNAL OF EUROPEAN CRIMINAL LAW 539-542 (2021); BRIAN SUNOHARA, *Credibility vs. Reliability* (Rogers Partners, Oct. 21, 2019), *available at:* https://www.rogerspartners.com/credibility-vs-reliability/#:~:text=Credibility%20has%20to%20do%20with,and%20recount%20events%20in%20issue (last visited 22 March 2025); SHANKAR LAW OFFICE, *Credibility v. Reliability in the Criminal Law Contex*, (Shankar Law Office, Sept. 24, 2020) *available at:* https://www.shankarlawoffice.com/post/credibility-v-reliability-in-the-criminal-law-context (last visited 22 March 2025). The relationship between credibility and reliability is subtle and can be best illustrated with an example; credibility has to do with a witness's veracity. Reliability has to do with the accuracy of the witness's testimony (his or her ability to accurately observe, recall and recount events). Credibility is not a proxy for reliability: a credible witness may give unreliable evidence. Credibility refers to the question of whether someone is telling the truth; reliability refers to whether an honest account of someone is accurate. The connection between reliability and validity is more intimate. For some, reliability is about accuracy (and they conflate

Contd…

approached in a, more or less, similar way. In practice, 'reliability' strongly correlates with the weight of evidence in a criminal trial.

But what stands is the fact that there is yet no such thing as 'a' harmonised European approach to evidence and the additional fact that codes in certain European states are rather general on principles of evidence law, such as reliability, exclusion and the use of experts. A recently concluded comparative research project on the use of AI in criminal justice, coordinated by the University of Luxembourg (*Katalin Ligeti* as principal investigator), reaches similar conclusions.[31] The inquisitorial/ adversarial divide in Europe is most poignant and largely responsible for these divergent national rules on admissibility and exclusion. Continental legal systems (*e.g.*, Belgium, the Netherlands or Germany), within the inquisitorial tradition, emphasise control on how evidence was obtained (and insist on the legality principle that implies that investigative measures need to be regulated); but they contain less detailed rules on admission, presentation and evaluation of evidence. The United Kingdom, like the US, relies heavily upon the jury system and the adversarial logic. It has detailed rules on admissibility with the judge in a gate-keeping role ensuring that the jury (trier of the fact) only sees admissible evidence that is produced in court. In all systems, included in the Luxembourg research study, there is a general tendency to admit AI evidence without too detailed scrutiny as to validity, reliability or credibility.

V

AI-Generated Evidence and Evidence Rules on Admissibility in Europe: The European Court of Human Rights

There is, in principle, a European consensus on the fair trial requirement that the prosecutor, judge, jury, and affected parties know that AI evidence is part of the evidence; yet national approaches vary regarding whether information on AI evidence is provided in the case file,[32] with situations, like in France, where the file may not contain information on the use of AI because it is considered not to be 'the evidence' but only to have generated 'leads' to evidence.

---

validity and reliability). Others, who distinguish between reliability and validity, hold that reliability is about consistency of evidence (is the account of a person consistent), while validity is about accuracy.

[31] K Ligeti, *Unpacking AI Evidence and (Re)-Defining Procedural Safeguards in Digital Investigations*, The Advent of AI: Reshaping Criminal Procedure, University of Luxembourg,(Nov. 7-8, 2024) *available at:* https://aiandcriminaljustice.uni.lu/2024/11/06/presentations-conference-crim_ai/ (last visited 12 Nov., 2024).

[32] The UK system requires indication if parts of the evidence were computer-generated or assisted; the Dutch system demands that reports introducing complex forensic evidence must indicate whether the evidence contains original or processed data.

Assuming that the defence is informed on the presence of AI evidence, there is equally consensus on the possibility of the defence to challenge admissibility. To do that, the defence needs to demonstrate that the AI output is not valid or reliable (or both) and/or has not been correctly applied to the case of the defendant.

At the European level, we note that both the European courts refuse to harmonise requirements of European evidence law.[33] They see it primarily as a task for the national judge to assess reliability. The ECtHR has nevertheless set out some requirements in terms of reliability and accuracy of evidence. For evidence to be considered reliable, there must be no doubt on its authenticity and integrity.[34] The ECtHR does pay attention to whether the defendant had the opportunity to challenge both the authenticity and the use of the contested (expert) evidence.

More problematic is the tendency of the ECtHR to only concentrate on manifest flaws, giving national courts a considerable margin of appreciation when assessing reliability and admissibility.[35] In a case of planted (by the police) evidence, the ECtHR identified three elements to make that assessment about the right to fair trial and reliability and admissibility: *one*, existence of procedural irregularities when collecting the evidence and vagueness concerning the actual circumstances of the police intervention; *two*, possible inadequate judicial and other remedies; and *three*, the existence of corroborating evidence.[36] Usually, when there is no manifest flaw, the assessment is more superficial. The ECtHR is then satisfied (and finds no violation of

---

[33] For a short summary of European case law, *see:* A Sachoulidou, *The Court of Justice in Staatsanwaltschaft Berlin v. M.N. (EncroChat)*,31(4) MAASTRICHT JOURNAL OF EUROPEAN AND COMPARATIVE LAW, 510-20 (2024); From cross-border, data-driven police investigations to evidence admissibility. *See* also: *Supra* note 13; A Završnik, *Criminal Justice, Artificial Intelligence Systems, and Human Rights*, 20 ERA FORUM 567-583 (2020); A Sachoulidou, *Going Beyond the "Common Suspects": To Be Presumed Innocent in the Era of Algorithms, Big Data and Artificial Intel*ligence, ARTIFICIAL INTELLIGENCE AND LAW (2023); F Palmiotto, *The Black Box on Trial: The Impact of Algorithmic Opacity on Fair Trial Rights in Criminal Proceedings* in ALGORITHMIC GOVERNANCE AND GOVERNANCE OF ALGORITHMS: LEGAL AND ETHICAL CHALLENGES 49-70 (Martin Ebers and Marta Cantero-Gamito, eds. 2020).

[34] K Quezada-Tavarez, P Vogiatzoglou & S Royer, *Legal Challenges in Bringing AI Evidence to the Criminal Courtroom,* 12 NEW JOURNAL OF EUROPEAN CRIMINAL LAW 531-551 (2021).

[35] *Id* at 531, 541. *See* also, *Khodorkovskiy and Lebedev* v. *Russia,* Applications 11082/06 and 13772/05, 72.

[36] *Kobiashvili* v. *Georgia,* Application no. 36416/06 (ECtHR, 14 March 2019). The three requirements were not met. a) Regarding the first requirement: no justification was given for a stop and search, hence the requirement of reasonable suspicion in domestic law could not be assessed. Also, it was not possible for the Court to understand the precise circumstances of the police intervention. The Court concludes that the way the search was carried out casts doubt on the reliability and accuracy of the evidence obtained as a result (para 65, b) The remedy requirement was not fulfilled: there was both lack of data on the information that the police had used to intervene and lack of responsiveness of the national court (that did not sufficiently consider the different statements and the possibility that there had been police pressure regarding the witnesses in this case) (para 71). c) Finally, regarding the corroborating evidence requirement, no other strong evidence of other sources was available (para 72).

the right to a fair trial), when the defendant was able to challenge the authenticity and use of the unlawfully obtained evidence, and it refuses to see a violation of the right to fair trial, when evidence is collected through a violation of the right to privacy or any other human right.[37] This last position, famously introduced by the Court in *Khan* (2000), seemingly reduces evidence law to address concerns about reliability and effective remedy, but excludes concerns on procedural irregularities, privacy violations, the rule of law and the legality principle.[38]

The result is that, contrary to the US legal system with its Federal Rules on Evidence and supporting case law, there is no such thing as a harmonised ECtHR-approach to reliability. All boils down to a case-by-case assessment and not deferring to strong statements about evidence law.

Still, when it comes to the general safeguards of fair trials, there is some certainty, as the ECtHR seems to insist, that the key question to be addressed remains the same: was the overall fairness of the proceedings ensured through the lens of procedural and institutional safeguards and the fundamental principles of a fair trial inherent in article 6 of the ECHR?[39] A dynamic interpretation of these fundamental

---

[37] *Khan* v. *The United Kingdom,* Application no 35394/97 (ECtHR, 4 October 2000 Final), paras 25-28 and 34-37; *Dragos Ioan Rusu* v. *Romania*, Application no 22767/08 (ECtHR, 31 January 2018 Final) para 49; *El Haski* v. *Belgium*, Application no 649/08 (ECtHR, 25 September 2012) para 83; *Jalloh* v. *Germany*, Application no 54810/00 (ECtHR, 11 July 2006) para 96.

[38] Recourse to devices to collect data without a legal basis in domestic law might violate Article 8 ECHR (requiring a legal basis and proportionality), but is not of a nature to exclude this illegally obtained data under Article 6 ECHR. K Quezada-Tavarez, P Vogiatzoglou & S Royer, *Legal Challenges in Bringing AI Evidence to the Criminal Courtroom*, 12(4) NEW JOURNAL OF EUROPEAN CRIMINAL LAW 531, 541, (2021), rightly ponder whether, 'given the increasingly dominant use of personal data in criminal cases, it would be worth examining whether the ECtHR should reconsider the relation between Article 6 and Article 8 ECHR'. This reconsideration has not taken part.

[39] *Yüksel Yalçinkaya* v. *Türkiye*, Application no 15669/20 (ECtHR, 26 September 2023) paras 311-313 ('(…) *Turning to the facts of the present case, it has already been established, in the context of the examination conducted under Article 7, that the applicant's conviction for membership of an armed terrorist organisation rested decisively on the finding that he had used the ByLock application, which finding was primarily based on the data obtained by the MİT, the remaining evidence serving only as a source of corroboration. In these circumstances, the quality of the evidence in question and the applicant's ability to effectively challenge it in proceedings that complied with the guarantees of Article 6 § 1 were all the more important. These matters will accordingly be examined in turn. However, before proceeding with this examination, the Court wishes to clarify whether the specific nature of the evidence at issue, that is encrypted electronic data stored at the server of an Internet-based communication application, requires it to adapt the application of the relevant guarantees under Article 6 § 1 in any way* (…) *The Court acknowledges that electronic evidence has become ubiquitous in criminal trials in view of the increased digitalisation of all aspects of life. It notes more pertinently, and without prejudice to its subsequent examination in the present case, that recourse to electronic evidence attesting that an individual is using an encrypted messaging system which had been specially designed for and used exclusively by a criminal organisation in the internal communications of that organisation, can be very important in the fight against organised crime* (…) *It also notes that electronic evidence differs in many respects from traditional*

Contd…

principles could, as case law of the ECtHR suggests, strengthen the position of the defence. For example, the principle of equality of arms can mean that the prosecution must present the defence with all relevant materials, including, not only evidence directly linked to the facts of the case but also, other materials that might relate to admissibility, reliability and completeness of such evidence.[40] Importantly, even though the interplay of various interests and their balancing could demand withholding some evidential materials from the defence, as well as, require that the defence give reasons for accessing such materials, any access-limitations should be strictly exceptional.[41] Despite the above useful insights, gained *via* case law, into how to interpret the fair-trial-basics, uncertainties appear to still engulf the overall fairness of the proceedings. If a specific AI-implementation does not qualify as (expert) evidence, but was nevertheless used and relied upon by law enforcement

---

forms of evidence, including as regards its nature and the special technologies required for its collection, securing, processing and analysis. Most significantly, it raises distinct reliability issues as it is inherently more prone to destruction, damage, alteration or manipulation. The Court also reiterates that the use of untested electronic evidence in criminal proceedings may involve particular difficulties for the judiciary as the nature of the procedure and technology applied to the collection of such evidence is complex and may therefore diminish the ability of national judges to establish its authenticity, accuracy and integrity (…) *Moreover, the handling of electronic evidence, particularly where it concerns data that are encrypted and/or vast in volume or scope, may present the law enforcement and judicial authorities with serious practical and procedural challenges at both the investigation and trial stages* (…) *That said, these factors do not in the abstract call for the safeguards under Article 6 § 1 to be applied differently, be it more strictly or more leniently. The Court's main concern is, once again, to assess whether the overall fairness of the proceedings was ensured through the lens of the procedural and institutional safeguards and the fundamental principles of a fair trial inherent in Article 6 of the Convention* (…)').

[40] *Matanovic* v. *Croatia*, Application no 2742/12 (ECtHR, 4 April 2017; final 4 July 2017) para 161 ('(…) *The Court notes that an issue with regard to access to evidence may arise under Article 6 in so far as the evidence at issue is relevant for the applicant's case, specifically if it had an important bearing on the charges held against the applicant* (…) *This is the case if the evidence was used and relied upon for the determination of the applicant's guilt or it contained such particulars which could have enabled the applicant to exonerate her- or himself or have his or her sentence reduced with regard to the charges held against him or her* (…) *It should be also noted that the relevant evidence in this context is not only evidence directly relevant to the facts of the case but also other evidence that might relate to the admissibility, reliability and completeness of the former* (…)').

[41] *Id*, paras 157, 152 ('(…) *Failure to disclose to the defence material evidence which contains such particulars as could enable the accused to exonerate him- or herself or have his or her sentence reduced would constitute a refusal of the facilities necessary for the preparation of the defence, and therefore a violation of the right guaranteed in Article 6 of the Convention* (…) *The accused may, however, be expected to give specific reasons for his request* (…) *and the domestic courts are entitled to examine the validity of these reasons* (…) *the entitlement to disclosure of relevant evidence is not an absolute right. In any criminal proceedings there may be competing interests, such as national security or the need to protect witnesses at risk of reprisals or to keep secret police methods of investigation of crime, which must be weighed against the rights of the accused* (…) *In some cases it may be necessary to withhold certain evidence from the defence so as to preserve the fundamental rights of another individual or to safeguard an important public interest. However, only such measures restricting the rights of the defence which are strictly necessary are permissible under Article 6 and 1* (…) *Moreover, in order to ensure that the accused receives a fair trial, any difficulties caused to the defence by a limitation on its rights must be sufficiently counterbalanced by the procedures followed by the judicial authorities* (…)').

in an opaque and secret fashion that the defence was unaware of, what reliability-related standards (if any) can apply?[42] The same applies for contestability, where AI tools are used, not as expert evidence but, to search for evidence, what contestability-standards can enable the defence to effectively challenge materials that may affect the probative value of evidence used against it?[43] And what about equal footing, in the context of finding excludable materials, in cases where advanced AI is used by the prosecution, yet the defence enjoys neither such advanced technologies (hence, probably) nor the right to access the results of the investigation and the secret or even unlawful way in which AI was deployed by state authorities?[44]

VI

AI-Generated Evidence and Evidence Rules on Admissibility in Europe: The Court of Justice of the European Union

The CJEU's position on evidence law echoes the ECtHR but has its specific EU law terminology and refers to specific EU law principles. Important is that, in the mindset of this Court, there is less 'hostility' towards exclusion of illegally obtained evidence. Thus, in Petar Dzivev and others, the CJEU essentially states that, while member states must ensure an effective fight against violations of EU laws (for instance, on VAT, Value Added Tax), this does not prevent the exclusion of illegally obtained evidence by national courts within domestic law frameworks, even if it is the only evidence on the basis of which the accused can be convicted.[45] In a way, the judgement can be read as a not so small adjustment of the ECtHR's Khan doctrine, particularly when the CJEU in *Dzivev* adds that European law does not dispense

---

[42] *Supra* note 13 at 208 ('(…) *possibly due to the interests of ongoing or future investigations, the use of the tools or the specific way they were employed may not always be revealed to the defence or the judge in the casefile* (…) *Moreover, (indirect) access for the defence to these tools may be necessary in order to adequately search data for possible exculpatory material* (…) *Finally, where the use of such tools is known, the use of AI tools for assisting in the analysis of data may not always be considered by law or courts as expert evidence, for which standards of scientific reliability and explication of method apply. For example, in Dutch literature and case law, it is often argued that the use of Hansken is not a form of expert evidence for which the standards of the ECtHR apply, as the tool is not used to produce an expert report in a formal sense, but rather to find evidence that can then be independently evaluated* (…) *The increasing complexity of the above mentioned tools, however, means they are likely to exert an increasing influence on an investigation or trial (the use of these AI tools might tilt the course of an investigation, influence the interpretation of materials later used as evidence, and involve violations of the right to private life or the right to the protection of personal data), and without information regarding their functioning or the way these tools have been employed, it is likely to be extremely difficult for both lawyers and judges to assess whether the investigation has been carried out in both a reliable as well as a lawful way* (…)').

[43] J Van Banning, THE AI ACT AS A SAFEGUARD FOR EQUALITY OF ARMS IN DATA-DRIVEN CRIMINAL INVESTIGATIONS, 212(2025).

[44] *Id.*

[45] *Petar Dzivev and others*, Case-310/16, CJEU (17 January 2019).

national courts from the necessary observance of the principle of legality and the rule of law; all the more so since the rule of law, as is clear from article 2 TEU (Treaty on European Union),[46] is one of the most important values on which the Union rests and privacy limitations need to be based on an explicit legal basis.

It is for national law alone to determine the rules on admissibility and evaluation of obtained information and evidence. In the absence of EU rules on that matter, it is, in accordance with the principle of procedural autonomy, for the national legal order of each member state to establish procedural rules for actions intended to safeguard the rights that individuals derive from EU law,[47] however, the important EU principles, such as the principle of equivalence and the principle of effectiveness, need to be respected.[48] It is for the competent national courts in criminal proceedings to ensure that these principles are safeguarded.

The CJEU struggled with its own cautious attitude on evidence law in *La Quadrature du Net and Others* (2020).[49] What happens when domestic courts are faced with evidence collected with data retention based on an EU law that has been declared invalid because of fundamental rights by the CJEU (in 2014 in its landmark decision *Digital Rights Ireland and Seitlinger*)?[50] What are the consequences of potential unlawful retention of or access to traffic and location data that are used as evidence in criminal proceedings?[51] The CJEU repeated its stance on evidence law (a national matter corrected by EU principles when there are rights at stake derived from EU law) and clarified that it does not follow from the principle of effectiveness that unlawfully obtained information and evidence used in proceedings against a person suspected of the commission of criminal offences needs to be prohibited as such.[52] Yet, the exclusion can be an option chosen by the national court when this is demanded for by

---

46 Treaty on European Union (2012) OJ C 326/13, Article 2.

47 *La Quadrature du Net and Others*, Joined Cases C-511/18, C-512/18 and C-520/18, CJEU (6 October 2020).

48 Under the principle of equivalence, domestic procedural rules intended to safeguard rights that individuals derive from EU law cannot be less favourable than rules governing similar domestic situations; and, under the principle of effectiveness, they do not render impossible in practice or excessively difficult the exercise of rights conferred by EU law.

49 *Supra* note 47.

50 *Digital Rights Ireland and Seitlinger a.o*, Joined Cases C-293/12 and C-594/12, CJEU (8 April 2014) para 223 with references to: *Târşia*, C-69/14, CJEU (6 October 2015); *XC and Others*, C-234/17, CJEU (24 October 2018); and *Deutsche Umwelthilfe*, C-752/18, CJEU (19 December 2019) para 33. *See*: A Juszczak & E Sason, *Recalibrating Data Retention in the EU*, 4 EUCRIM 238 (2020).

51 This question was raised by the Belgian *Cour Constitutionelle* that sought to address that issue by inquiring whether it may maintain the effects -at least temporarily- of the national law on data retention under review, even if the CEU has found that it does not comply with EU law.

52 Other means may too serve the purpose of not prejudicing that person unduly by using unlawfully obtained material, such as national rules and practices governing the assessment and weighting of such information and evidence or the consideration whether such material is determining the sentence.

the adversarial principle as part of the right to a fair trial .[53] This would be the case when the national court finds that a party is not able to effectively comment on evidence which pertains to a field of which the judges have no knowledge, and which is likely to have a preponderant influence on the findings of fact.[54]

When introducing these fair trial requirements, the CJEU refers to one of its previous judgements *Steffensen*[55] on food safety. *Steffensen* refers to ECtHR's *Mantovanelli* v. *France* on liability in a case of medical maltreatment.[56] Both cases are relevant for our contribution. They deal with results of very technical expert analysis and reports (on the quality of meat and the acceptable amount of a drug) that went beyond the technical knowledge of the national court. Admission as evidence of these expert findings entailed, in the view of the respective European Courts, a risk of infringement of the adversarial principle, given that, in the *Steffensen*-case, the party was not given a right to request a second opinion in violation of EU law, while, in *Mantovanelli,* the party was entirely excluded from the preparation of the expert report. Thus, although the administrative courts were not legally bound by the expert findings, these were likely to have a preponderant influence on the assessment of the facts. Juszczak and Sason argue that the CJEU transposed in a questionable way this earlier case-law to the field of criminal law and the matter of data retention,[57] but it can nevertheless be asserted that for the CJEU because of the requirements of the principle of effectiveness, national criminal courts should disregard information and evidence obtained in the context of criminal proceedings where the defendants are not in a position to comment effectively on information or evidence that pertain to a field of which the judges have no knowledge and that are likely to have a preponderant influence on the findings of facts.[58]

In her presentation of the national reports that are part of the comparative research discussed *above*, Ligeti[59] notes a tendency very similar to what we found in our previous work on the US to move from blind trust to real scrutiny by the courts. Based on analysis of case law from the European and domestic courts, the research finds that European domestic courts increasingly take a more defence protective stance by requiring that AI evidence is supported by other proof and that conclusions are based on and controlled by human judgement. Also, there are novel approaches to recognise (new) defence rights. We will come back to these new rights in a later section.

---

[53] *Supra* note 47 at 226.

[54] *Supra* note 47 at 227.

[55] *Joachim Steffensen*, Case C-276/01, CJEU (10 April 2003).

[56] *Mantovelli* v. *France*, Application No 21497/93, ECtHR (18 March 1997).

[57] A Juszczak & E Sason, *Recalibrating Data Retention in the EU*, 4 EUCRIM 238 (2020

[58] For a brief, but to the point critical summary of the European case law on evidence, *Supra* note 54. S*ee* also: A Sachoulidou, *The Court of Justice in Staatsanwaltschaft Berlin v. M.N. (EncroChat)*,31(4) MAASTRICHT JOURNAL OF EUROPEAN AND COMPARATIVE LAW, 510-20 (2024).

[59] *Supra* note 31.

VII

Expert Evidence Demands in Europe: Lack of Harmonisation

Academics have identified a general tendency in Europe to admit AI evidence without too detailed scrutiny as to validity, reliability or credibility.[60] Problems with the determination of evidence reliability and authenticity are directly related to the question of expert knowledge, since this determination in the case of AI evidence requires specialised fact-finding.

The presence of experts in court is a global phenomenon. The prosecutor is obliged to disclose both inculpatory and exculpatory evidence to the defence, including if the evidence contains forensic reports. In our previous work on the US legal system, we found detailed rules on the use of experts in court and the control by the judge over the scientific nature of the experts and their methods. These rules are contained in the Federal Rule of Evidence 702 and further elaborated in settled case law (following *Frye*,[61] *Kelly*[62] and *Daubert*).[63] In general, it can be said that Europe has set out some minimum (US-like) requirements focusing on: competence (that can be actual expertise, similar to the US qualification on the basis of, among others, actual experience); independence[64] or assistance (same as in the US, the expert must provide for data that will assist the court); or reliability in terms of, among others, data-quality and quantity, accuracy of methodologies adopted, reviewability or acceptance of a given practice in the relevant scientific area[65] (elements similar to the *Daubert*'s minimums and other US requirements, *e.g.*, to use reliable methods and reliably apply them to a given case).

Decaigny, a legal scholar well-versed in defence rights from a comparative perspective, paints a subtle picture of expert testimony in inquisitorial or in adversary systems and convincingly sees strong points in *both*.[66] In an adversarial criminal trial, experts, much like witnesses, will be proposed or called in by a party. The accused's fair trial rights, also regarding expert testimony, are protected with a focus on the equality of arms requirement, typically asserted in open court at the

---

60 *Supra* note 31.

61 *Frye* v. *United States*, 293 F 1013 (DC Cir 1923).

62 *People* v. *Kelly*, 17 Cal 3d 24 (Cal 1976).

63 *Daubert* v. *Merrell Dow Pharmaceuticals Inc*, 509 US 579 (1993).

64 On independence of experts, *see*: *Yüksel Yalçinkaya* v. *Türkiye*, Application no 15669/20 (ECtHR, 26 September 2023) paras 332-333. *See* also a comparative analysis (mainly of Federal Rule of Evidence 706 and the French scheme) in: W J Travis, *Impartial Expert Testimony Under the Federal Rules of Evidence: A French Perspective*, 8(3) *International Lawyer* 492 (1974).

65 R Glover, MURPHY ON EVIDENCE, 612-622 (2017). On evidence-related law and procedural safeguards, *see* also: S Gless, *Evidence and Proof* in ELGAR ENCYCLOPEDIA OF CRIME AND CRIMINAL JUSTICE (P Caeiro et al, eds. 2024); K Ligeti, AI EVIDENCE AND CRIMINAL PROCEEDINGS (Hart Publishing forthcoming 2025).

66 T Decaigny, *Inquisitorial and Adversarial Expert Examinations in the Case Law of the European Court of Human Rights*, 5(2) NEW JOURNALS OF EUROPEAN CRIMINAL LAW 164-166 (2014).

trial stage. Each party can make use of its own expert, and their findings are tested in court. The resulting confrontational approach in adversarial cases has the advantage of a full debate in open court but also has disadvantages.[67]

In a continental or inquisitorial process, on the other hand, there is a common, court-appointed expert, leading expert examinations in which a party could participate.[68] The advantage that Decaigny sees is forum separation: formally all fact-finding occurs at the trial stage, but important aspects of criminal investigation, including expert examination, often takes place in the pre-trial:

> (…) *Pragmatically conducting part of the fact-finding in the expert examination, prior to the final expert report and the debate upon that report, allows a consideration of the expert examination as a forum-separation. Subsequently, fair trial rights can be guaranteed to the largest extent by allowing the defence to intervene at the moment of the de facto fact-finding, i.e., during the expert examination*(…).[69]

Moreover, the inquisitorial possibility of defence participation in the expert examination offers the opportunity of a more constructive discussion, led by the court-appointed expert (and possibly joined by the partisan experts).

The foregoing is a soft defence of the inquisitorial way to handle experts. Useful as it is, it is nevertheless undeniable that inquisitorial systems struggle with expert evidence due to historical reasons. For instance, the Belgian Code on Criminal Procedure[70] has still no systematic chapter on evidence testimony. Other continental codes fare better, but all fundamentally share a firm belief in 'the truth' controlled by the judge and are horrified by the idea that all parties come with 'their experts' and confront 'their truth' with others' truth.[71] The result is a system of expert evidence, where the parties can ask for it, but the judge (either the investigatory judge or the trial judge) must take the decision to appoint one.[72] Another result is an unsatisfactory mechanism to challenge the results of expert testimony and have them vetted in a meaningful way (*e.g.*, confrontation rights, second opinions or

---

[67] *Supra* note 66 at 165.

[68] *Id* at 164.

[69] *Id* at 166.

[70] Belgian Code on Criminal Procedure, 1978.

[71] *Supra* note 66 at 165.

[72] The requirement to be able to challenge evidence is now recognised in inquisitorial systems, because of the defence rights contained in Article 6(3d) ECHR and *Supra* note 66. Concerning expert evidence, this is clear from *CB* v. *Austria*, Application no 30465/06, ECtHR, 4 April 2013, para 43, where the ECtHR insisted on the possibility to challenge an inquisitorial court-appointed expert. The same court also emphasised that experts may not be eyewitnesses, but they are nevertheless witnesses within the meaning of Article 6 (3d) ECHR implying that the defence must be able to study and challenge both the expert report and the person who prepared it through questioning. *See*: *Mirilashvili* v. *Russia*, Application no 6293/04 (ECtHR, 11 December 2008, final 5 June 2009) paras 158-160; *Khodorkovskiy and Lebedev* v. *Russia*, Applications 11082/06 and 13772/05, 711.

access to AI systems).[73] To Gless, the European inquisitorial systems insufficiently address the 'science dilemma', the problem of how to determine in a concrete case: (a) what intellectual enterprise that might yield expert testimony is a science that is *rationally pertinent* to the case; (b) who is a *scientist* capable of using her/his knowledge in a manner that satisfies the standard of epistemic appraisal and the attendant level of confidence.[74]

This problem, touched upon by Gless, is particularly relevant. When discussing the reliability of evidence *above*, we observed that questions of reliability arise when human experts operating AI tools produce different results. Experts called upon in court to speak out on reliability of evidence can equally jeopardise reliability of AI evidence. There is low reliability between digital forensics experts examining the same evidence file with the same contextual information, when their observations, interpretations and conclusions differ.

All criminal justice systems included in the Luxembourg research lacked standardised tests for forensic evidence. Often it was found that the defence lacked the means to pay for forensic experts.[75] The finding is not neutral and corrects Gless' thesis that adversarial systems have a superior level of detail (on the 'science dilemma'), but have a cost problem for the defence, putting pressure on the idea of equality of arms.

Interesting is the proposal by Gless and others[76] for the following starting points to be respected when AI evidence is presented in court; and this, to mitigate the science bias problem with AI in court and to address the cost problem that arises when experts are needed:

a. Experts should be used where a trier of fact is not epistemically competent to assess a scientific claim.
b. Admissibility of expert testimony requires standardised tests as to what scientific knowledge is, and what its limits are.
c. Results of expert testimony need to undergo meaningful vetting of its result (confrontation rights, second opinions, access to models of devices / AI systems etc).
d. Legal defence funds must ensure that expert testimony respects 'equality of arms'.

---

[73] *Supra* note 8; P W Grimm, M R Grossman & G V Cormack, *Artificial Intelligence as Evidence*, 19(1) NORTHWESTERN JOURNAL OF TECHNOLOGY AND INTELLECTUAL PROPERTY 9 (2021).

[74] S Gless, *Rules on Expert Testimony based on a Comparative Perspective on Device Evidence*, The Advent of AI: Reshaping Criminal Procedure, University of Luxembourg, (7-8 November 2024) *available at:* https://aiandcriminaljustice.uni.lu/2024/11/06/presentations-conference-crim_ai/ (last visited 12 Nov., 2024).

[75] *Supra* note 31.

[76] E Silverman, J Arnold & S Gless, *Robot Testimony? A Taxonomy and Standardized Approach to the Use of Evaluative Data in Criminal Proceedings* in S Gless & H Whalen-Bridge (eds), HUMAN–ROBOT INTERACTION IN LAW AND ITS NARRATIVES 167 (S Gless & H Whalen-Bridge, ed. 2024); B Garrett & C Rudin, *The Right to a Glass Box: Rethinking the Use of Artificial Intelligence in Criminal Justice*, 109 CORNELL LAW REVIEW 561-626 (2024); *Supra* note 73.

These starting points (that include epistemic competence, standardised tests, meaningful contesting and adequate legal defence funding) can be seen as a minimum defence-shielding package to tackle expert-evidence-related challenges posed by the introduction of AI in criminal trials and ensure, in practice, an equal footing, especially, where it might not be clear whether the defence enjoys the right to bring its own expert to participate in investigations and, thus, effectively contest evidence used against it.[77] With more and more actors entering the scene, including commercial actors with financial interests, technical experts being increasingly needed to enlighten judges and explain what might be humanly unintelligible, and with expert evidence escaping and being evaluated outside the courtroom, fairness in criminal trials may need to be reconstructed to ensure transparency and reliability for the benefit of the defence.[78]

VIII

Fair Trial Demands in Europe: Robust Defence Rights-Scheme, but Unstable Proof-Thresholds

The European fair trials scheme (under Article 6 of the ECHR), in general, provides for some minimums, resembling the US general requirements of due process.[79] These includes, the defence's genuine opportunity to present its arguments and challenge allegations against it (including the chance to contest admissibility and reliability of evidence),[80] equality of arms and adversarial proceedings,[81] the

---

[77] *Supra* note 13, with further references to case law on the right of the defence to comment on and study the evidence and the right to call its own expert.

[78] Palmiotto F, *Artificial Intelligence and the Transformation of Criminal Trials*, 31 EUR J CRIME CRIM L & CRIM JUST 155 (2023); In her PhD thesis, Palmiotto analyses the transformative effect of AI in criminal trials and argues for the need for legal reforms to ensure access to information and experts for the defence.

[79] *Id*; R Kailas Bharati, *AI and the Right to Fair Trial: Legal Standards and Technological Advances*, 5(7) INTERNATIONAL JOURNAL OF SCIENCE ACADEMIC RESEARCH 7865 (2024); K Terzidou, *The Use of Artificial Intelligence in the Judiciary and Its Compliance with the Right to a Fair Trial*, 31 JOURNAL OF JUDICIAL ADMINISTRATION 154 (2022).

[80] On the defence's opportunity to contest the integrity and reliability of evidence, *supra* note 39. *See* also: *Jalloh* v. *Germany*, Application no 54810/00 (ECtHR, 11 July 2006) para 96; *Allan* v. *United Kingdom*, Application no 48539/99 (ECtHR, 5 February 2003 Final) para 43; *Dusko Ivanovski* v. *the former Yugoslav Republic of Macedonia*, Application no 10718/05 (ECtHR, 24 July 2014 Final) para 43; *Prade* v. *Germany*, Application no 7215/10 (ECtHR, 3 June 2016 Final) para 34; *Khan* v. *the United Kingdom*, Application no 35394/97 (ECtHR, 4 October 2000 Final) para 38.

[81] Council of Europe, *Guide on Article 6 of the European Convention on Human Rights – Right to a Fair Trial (Criminal Limb)*, (updated on 29 February 2024) 35, *available at:* https://ks.echr.coe.int/documents/d/echr-ks/guide_art_6_criminal_eng (last visited 24 Oct., 2024).

presumption of innocence[82] or the requirement of proof beyond reasonable doubt.[83] Stoykova identifies no less than 10 guarantees of fairness and considers them as a source of European evidence rules.[84]

Case law on article 6 of ECHR is impressive and the scope of its requirements is expanding. But the overall picture, in Europe, remains one of fragmentation and pluralism, partly because the ECtHR is not always very activist and partly due to the presence of both inquisitorial and adversarial systems in Europe.[85] However, given national divergencies, Europe seems to suffer weaknesses in terms of harmonisation, *e.g.*, with different degrees of probability, when it comes to proof beyond reasonable doubt.[86]

We have already highlighted Ligeti's finding[87] of the tendency to move from blind trust to real scrutiny by the courts. Interesting is her finding that, at least in some legal systems, there are novel approaches to recognise (new) defence rights:

a. Right to information on the AI tool.
b. Right to access to the full collection of data.
c. Right to access the AI tool.
d. Right to explanation of forensic methods and results.
e. Right to have digital forensic assistance.

The first right can be partly based on the 2024 AI Act (see *below*), but all others are elaborations made by courts based on, *e.g.*, the notion of effective defence rights. Ligeti's list gives a taste of the real battle going on in courts nowadays.[88] With regard

---

[82] European Convention on Human Rights, 1950, art 6, para 2; Council of Europe, *Guide on Article 6 of the European Convention on Human Rights*, 2006. It is noted that the European Union's approach to fair trials is based on the ECHR's provisions. *See*: Charter of Fundamental Rights, (14 December 2007) OJ C303/2007, arts 47-48. On due process and fair trials, see, in general: C Momsen & M Willumat, *Due Process and Fair Trial* in ELGAR ENCYCLOPEDIA OF CRIME AND CRIMINAL JUSTICE 102 (P. Caeiro et al ed.).

[83] Council of Europe, *Guide on Article 6 of the European Convention on Human Rights*, 2006. *See* also: H Jellema, *Reasonable Doubt, Robust Evidential Probability and the Unknown*, 18 CRIMINAL LAW AND PHILOSOPHY 451(2024), discussing proof of guilt beyond a reasonable doubt and recommending a novel way to combine the Bayesian and explanation-based accounts of rational legal proof.

[84] R Stoykova, *The Right to a Fair Trial as a Conceptual Framework for Digital Evidence Rules in Criminal Investigations*, 49 COMPUTER LAW AND SECURITY REVIEW 105801 (2023).

[85] P De Hert, *Legal Procedures at the International Criminal Court. A Comparative Law Analysis of Procedural Basic Rights* in R Haveman, SUPRANATIONAL CRIMINAL LAW: A SYSTEM SUI GENERIS, 4.6 (O Kavran & J Nicholls ed., 2003).

[86] C Bicknell, *Uncertain Certainty? Making Sense of the European Court of Human Rights' Standard of Proof*, INTERNATIONAL HUMAN RIGHTS LAW REVIEW (2019), *available at:* http://hdl.handle.net/10871/38992 (last visited 24 Oct., 2024).

[87] *Supra* note 31.

[88] The Dutch Marengo appeal judgement, Court of Appeal Amsterdam, 25 October 2024, ECLI:NL: GHAMS:2024:2975, with regard to the use by law enforcement authorities of the Hansken AI system is a good example. *See* also the Supreme Court (28 June 2022,

Contd…

to the second right, it is still unclear why the ECtHR refuses to recognise such a right and limits the duty to search for exculpatory evidence only to data identified as relevant by the prosecution and not to the whole data set.[89]

From the perspective of effective defence rights and equality of rights, it would be an important step forward to have these new approaches (to be detected in some legal systems) formally recognised and rolled out in the European legal systems (see *below*). The fifth proposed right aligns well with the starting points for expert testimony in AI recommended by Gless (*above*, previous section), in particular with her fourth proposal to create legal defence funds. Note that there are other proposals in literature, such as the right to procedural accuracy proposed by Stoykova[90] and the many ideas proposed by Francesca Palmiotto.[91]

A remark can be made on Europe's (the EU's) approach to trade secrecy, protected under the Trade Secrets Directive.[92] The protection of trade secrets against unlawful acquisition, use and disclosure, under this Directive, resembles that of the US, in terms of definitions, prohibitions or exemptions.[93] Regarding disclosures of trade secrets in criminal proceedings, most European national schemes provide for procedural rules, offering the defence the opportunity to access trade secrets under confidentiality-safeguards (*e.g.*, closed hearings).[94] Although such a conditional access can also be granted in the US *via* protective orders or other legal instruments, it is often denied, in light of developments in case law, favouring the trade secrecy

---

ECLI:NL:HR:2022:900; 23 May 2023, ECLI:NL:HR:2023:743-746). On this first AI case in the Netherlands, R. Robroek, 'Algorithmic criminal procedure in the Netherlands' (The Advent of AI: Reshaping Criminal Procedure, University of Luxembourg, 7-8 November 2024) *available at:* https://aiandcriminaljustice.uni.lu/2024/11/06/presentations-conference-crim_ai/ (last visited 12 Nov., 2024).

89 *Sigurður Einarsson and others* v. *Iceland*, Application no 39757/15 (ECtHR, 4 June 2019, final, 4 September 2019) para 90; *Rook* v. Germany, Application no 1586/15 (ECtHR, 25 July 2019, final, 25 October 2019) para 67, 72; *See:* R Stoykova, *The Right to a Fair Trial as a Conceptual Framework for Digital Evidence Rules in Criminal Investigations*, 49 COMPUTER LAW AND SECURITY REVIEW 105801 (2023).

90 R Stoykova, *A New Right to Procedural Accuracy: A Governance Model for Digital Evidence in Criminal Proceedings*, 55 COMPUTER LAW AND SECURITY REVIEW 106040 (2024).

91 F Palmiotto, *The Black Box on Trial: The Impact of Algorithmic Opacity on Fair Trial Rights in Criminal Proceedings* in ALGORITHMIC GOVERNANCE AND GOVERNANCE OF ALGORITHMS: LEGAL AND ETHICAL CHALLENGES 49-70 (Martin Ebers and Marta Cantero-Gamito, eds. 2020), discussing legal challenges posed by algorithmic opacity in criminal proceedings and claiming that algorithms may contain miscodes, opacity in algorithms can keep such miscodes hidden, and, thus, algorithmic opacity can have a hostile impact on fair trial rights.

92 Trade Secrets Directive, 2016 (EU) 943 OJ L157/1.

93 *Supra* note 92 at arts 2-5.

94 European Union Intellectual Property Office, *The Baseline of Trade Secrets Litigation in the EU Member States* 14 (European Union Intellectual Property Office 2018), offering a state-by-state analysis, demonstrating that, in the vast majority of national systems, some type of minimum access (e.g., via hearings behind closed doors) is granted, where trade secrets are involved in criminal proceedings.

privilege.[95] This finding is confirmed by the research project led by Ligeti. [96] In her report, it is noted that courts in the US and Europe often deny defence requests to access the information required for an independent validation. This often happens in the US due to trade secrecy protection (of proprietary AI). In Europe, as seen, this ground for refusal is less invoked, but this does not seem to improve the situation from a defence perspective. Other justifications are used to deny access, for example, that such a discovery is not necessary for disposing fairly the action (in the Netherlands and the United Kingdom) or that it will incur unnecessary costs (in the United Kingdom).

IX

Equal Protection (non-discrimination) Demands in Europe: No Need to Show Discriminatory Intent

Concerning the prohibition of undue discrimination (protected under article 14 of the ECHR), Europe, identical as the US, prohibits unfair discrimination on the basis of protected grounds. Yet, the scrutiny-levels applied to ascertain undue discrimination are not as clearcut as the US' strict/intermediate/rational-scrutiny.[97] An important point that could make the European regime advantageous, compared with the US scheme on equal protection, is Europe's focus on the *discriminatory effect*[98] and the absence of the requirement to demonstrate *discriminatory intent or purpose*.[99] In the artificial intelligence context, this could favour the defence that would not need to demonstrate that an AI-provider or deployer, engaging in an unfairly discriminatory practice, actually had the intention to mistreat someone.[100]

---

[95] *See* our previous analysis of the right to confrontation and the trade secrecy privilege in: De Hert P and Bouchagiar G, *Law Enforcement Relevant Regulation* in ELGAR ENCYCLOPEDIA OF CRIME AND CRIMINAL JUSTICE 8 (P Caeiro et al,ed. 2024).

[96] *Supra* note 31.

[97] European Convention on Human Rights, 1950, Article 14, offers an indicative list of protected grounds. The same is the case with the EU, under Article 21 of the CFR.

[98] On direct and indirect discrimination, *see*: *Biao* v. *Denmark*, Application no 38590/10 (ECtHR, 24 May 2016 Final) paras 89, 103. On the EU's approach to direct/indirect discrimination, *see*: Council Directive 2000/43/EC of 29 June 2000 implementing the principle of equal treatment between persons irrespective of racial or ethnic origin (2000) OJ L180/22, art 2, para 2.

[99] European Union Agency for Fundamental Rights and Council of Europe, *Handbook on European Non-Discrimination Law* (2018). *See* also: *DH and Others* v. *the Czech Republic*, Application no 57325/00 (ECtHR, 13 November 2007), para 184; *Biao* v. *Denmark*, Application no 38590/10 (ECtHR, 24 May 2016 Final), paras 103.

[100] It is added that case law of the CJEU on race and sexual orientation has held that there is no requirement to demonstrate that there is an identifiable victim. See: *Centrum voor gelijkheid van kansen en voor racismebestrijding* v. *Firma Feryn NV*, Case C-54/07 (10 July 2008) para 40; *Asociația Accept* v. *Consiliul Național pentru Combaterea Discriminării*, Case C-54/07 (25 April 2013) para 62.

Contd…

X

Data Protection, the Legality Principle, Guarantees, and Oversight

The legality principle is as important in Europe as it is in the United States. But the ways in which it is applied by the two jurisdictions differ.[101] While drafting laws, the approach of the Europe to privacy, for instance, is more complete, but also more abstract and technologically neutral, compared with the US piece-meal and targeted technique.[102] We may recall that the EU 2016 Law Enforcement Directive (LED) is a *lex specialis* published together with the GDPR.[103] The LED's scope covers operational police work as well as the work of public prosecutors and investigative judges, criminal courts and the operation of prisons. It further covers the processing by 'competent authorities' in the law enforcement and the criminal justice sector.[104] It clarifies the legality principle for the processing of data by these actors and enumerates important processing principles (*e.g.*, data quality, minimisation, human oversight, purpose limitation or notification rights). It recognises (as important) data subject rights, like access and notification, and sets up a system of administrative oversight by national data protection authorities that can give guidance on the application of the LED in the AI context.[105] These authorities are an important 'newcomer' in the law enforcement world traditionally dominated by police, prosecutors and judges.

But this would hardly apply to the ECtHR-regime, considering the admissibility requirements under the European Convention of Human Rights, 1950, art 34.

[101] Above, we found that European Courts are not prone to exclude evidence on the ground of procedural irregularities or when a legal basis is lacking. *Petar Dzivev and Others* v. *Prokuratura na Republika Bulgaria* (Case C-310/16, Court of Justice of the European Union, 17 January 2019) ECLI:EU:C:2019:30, rightly adjusted a possible *laissez faire* attitude by recalling the importance of the legality principle and the rule of law. So, laws matter, and Europe is producing them.

[102] In the EU, the legal regime on personal data protection, as framed under the General Data Protection Regulation, 2016, opts for a rather tech-neutral approach. Moreover, the legal framework, established under the Law Enforcement Data Protection Directive, 2016, appears police-friendly, lacking enhanced checks and balances (e.g., the principle of transparency). For a critical discussion, including pros and cons of technological neutrality in lawmaking, *see*: P De Hert & G. Bouchagiar, *Visual and Biometric Surveillance in the EU. Saying 'No' to Mass Surveillance Practices?* 27(2) *Information Polity* 193(2022).

[103] Council Framework Decision 2008/977/JHA (2016) OJ L119/89 ('Law Enforcement Directive'), 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.

[104] For the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, including safeguarding against and prevention of threats to public security.

[105] *See:* EU AI Act, 2024, S.11.

The LED is silent on interference with the specific language of criminal and police-related law. It also lacks on integration of the jurisprudence of the ECtHR and the CJEU on surveillance investigative and police practices. The latter is even more serious, since there is a focus in the jurisprudence of the European courts not only on the language of fundamental rights, but also the language of personal data protection, as well as, where necessary, the language of criminal and police-related law. The latter language is characterised by, among other things, building hierarchical and judicial supervision either prior to the measure (*e.g.*, the police must ask for permission to use a certain AI system) or afterwards.[106] The European broad data protection approach may not expressly address particularities of AI-made decisions and content. There are many questions, therefore, which may not be easily answered: including, 'to what extent can state actors using AI explicate opaque and sophisticated decisions? How informed are citizens, as data subjects, concerning AI-decisions acted upon them? or How could the principle of data minimisation be respected in light of AI's disposition toward mass data processing?

Considering the European human rights law, the foregoing is, in our opinion, not enough. Europe's approach to the principle of legality (also known from and well-developed by case law of the ECtHR)[107] demands for the introduction of concrete laws[108] that are adequately accessible,[109] foreseeable[110] and detailed, expressly referring to whether and the way in which law enforcement or criminal justice authorities may consider AI-made decisions and contents. Remember the American scholar Garrett's proposal for regulation to safeguard in all criminal cases the right to glass box AI. This is the opposite of technological neutrality. In the AI-context, foreseeability, as addressed by case law of the ECtHR, could require that laws make specific reference to, *e.g.*, the categories of people liable to be subjected to an AI-related law or practice, the nature of the crimes, triggering AI-applications, limitations regarding duration of the AI-related law or practice (limiting AI-uses in specific phases of the criminal procedure), procedures of data processing *via* the use

---

[106] P De Hert & J Sajfert, *Variety, velocity, and volume of personal data in criminal investigations and proceedings: the limits drawn by the purpose limitation and data minimization principles in Directive (EU) 2016/680'*, in VALSAMIS MITSILEGAS, RESEARCH HANDBOOK ON EU CRIMINAL LAW 21 (Maria Bergström & Teresa Quintel, 2nd ed., 2024); C Jasserand, *Law enforcement access to personal data originally collected by private parties: Missing data subjects' safeguards in directive 2016/680?*, 34(1) COMPUTER LAW AND SECURITY REVIEW 154 (2018).

[107] P De Hert & G Malgieri, *One European Legal Framework for Surveillance: The ECtHR's Expanded Legality Testing Copied by the CJEU* in SURVEILLANCE AND PRIVACY IN THE DIGITAL AGE. EUROPEAN, TRANSATLANTIC AND GLOBAL PERSPECTIVES 255 (V Mitsilegas & N Vavoula eds., 2021).

[108] The legality assessment requires a national legal basis. *See*: *Huvig* v. *France*, Application no 11105/84 (ECtHR, 24 April 1990) paras 26-28; *Kruslin* v. *France*, Application no 11801/85 (ECtHR, 24 April 1990) paras 27-29.

[109] *Malone* v. *the United Kingdom*, Application no 8691/79 (ECtHR, 2 August 1984) para 66.

[110] On foreseeability, *see*, among others: *Huvig* v. *France*, Application no 11105/84 (ECtHR, 24 April 1990) para 29; *Kruslin* v. *France*, Application no 11801/85 (ECtHR, 24 April 1990) para 30; *Id* at 67, 70.

of AI, precautions on sharing data (limiting the group of people authorised to process relevant information), or the circumstances for data erasure (through setting out deletion-periods).

XI

The AI Act, the Legality Principle: Guarantees and Oversight

The AI Act, 2024, is a second piece of European law that essentially concerns our present theme.[111] This Act divides AI applications into risk groups. The *first category* is of AI applications with an unacceptable risk. These are prohibited according to article 5 of the Act. These include applications that create databases for face recognition, work on emotion recognition at work or school and applications that try to abuse vulnerabilities of certain groups of persons (*e.g.*, minors or the disabled). The AI Act also provides for an explicit prohibition on AI applications that predict how likely it is that a person will commit a crime. It will not be possible for police authorities using such an AI application to take preventive measures. The *second category* consists of AI applications with a high-risk profile, to which much attention is paid in the AI Act. These are applications that could negatively impact fundamental rights. Most systems of AI fall under this category. These systems must be capable of being effectively controlled by a natural person throughout their lifetime (article 14 of the AI Act).[112] A *third category* an AI application relates to transparency risks. These are, *e.g.*, applications that communicate directly with humans, such as chatbots. These applications are required, by article 50 of the AI Act, to inform the user that they are communicating with an AI application and not a natural person. The same article refers to AI applications that can produce or manipulate artificial images and sounds. These applications are also subject to the obligation to inform the user that they are artificially generated images and sounds.

In a previous work, we highlighted the relation between the LED and, what we have called the *newcomer*, in particular the added value of the Fundamental Rights Impact Assessment (FRIA), under article 27 of the AI Act; and a possible reading of article 13(1) of the AI Act together with article 86 of the AI Act, in the sense that these provisions could require deployers of the AI tools to provide deeper insights into

[111] Artificial Intelligence Act, OJ L2024/1689 (12 July 2024). 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.

[112] In other words, these applications require that a human being must always be able to check whether they function properly. They must also allow this person to overrule or even stop the AI application.

their functioning, not only disclosing information on the algorithms applied, but also details on the model's architecture and/or the data used for its training.[113]

This is, however, what the AI Act, even in the most constructive reading, can deliver. Built upon the regulatory choices of a risk-based and a product safety perspective,[114] this Act seems to emphasise risk[115] over the logic of the 'fundamental',[116] probably failing to adequately protect those who could be affected by AI. In Edwards' terminology,

> (t)*hose impacted by AI systems – sometimes thought of as end-users, data subjects or consumers – have no rights, and almost no role in the AI Act*.[117]

Despite previous discussions on the acute need to respect fundamental principles and rights,[118] the AI Act does not set out general principles applicable to *all* AI-uses.[119] The Act introduces important limitations on law enforcement's use of AI, but does not impose stringent obligations and clear bans on certain AI-applications, including systems that enable biometric mass surveillance and predictive policing systems.[120] Another prime example is the use of AI to assess the risk of engaging in criminal conduct, that is (in principle) prohibited, yet may be permitted, when aimed at, among others, supporting the human decision on someone's involvement

---

[113] P. De Hert, *al. el.*, , '*Evidence Law, Data Protection Standards and the Use of AI for Investigation and Prosecution*' in K Ligeti & N Visco AI EVIDENCE AND CRIMINAL PROCEEDINGS (Hart Publishing forthcoming).

[114] T Evas, *The EU Artificial Intelligence Act: Advancing Innovation for Trustworthy AI*, 1 AIRE 98 (2024); N A Smuha & V Hendrickx, *AI and the Administration of Justice: Taking 'Precedent Analysis' as a Use Case to Assess the Adequacy of the AI Act*, (KU Leuven, Oct 2, 2023), *available at:* https://www.law.kuleuven.be/ai-summer-school/blogpost/Blogposts/AI-administration-justice (last visited 24 Oct, 2024).

[115] On the lack of clear reviewable criteria supporting the AI Act's risk-classification, *see*: L Edwards, *Expert opinion: Regulating AI in Europe. Four problems and four solutions*, (Ada Lovelace Institute, March 31, 2022), *available at:* https://www.adalovelaceinstitute.org/report/regulating-ai-in-europe/ (last visited 24 Oct, 2024).

[116] For example, there is no impact assessment regarding fundamental rights to be conducted for all AI uses. *See*: Editorial, *There Are Holes in Europe's AI Act — and Researchers Can Help to Fill Them*, 625 NATURE 216 (2024).

[117] *Supra* note 115.

[118] See, among others: European Parliament, *Artificial intelligence in criminal law and its use by the police and judicial authorities in criminal matters* (European Parliament resolution of 6 October 2021 on artificial intelligence in criminal law and its use by the police and judicial authorities in criminal matters), P9 TA (2021)0405 (2020/2016(INI)) paras 2, 4.

[119] The European Parliament's proposal for an article (4 a) on general principles applicable to all AI systems was abandoned and, eventually, watered down and reduced to requirements for high-risk AI systems or obligations of providers/deployers of high-risk AI systems (see Amendments Adopted by the European Parliament on 14 June 2023 (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD)), art 4 a).

[120] The lack of clear bans was stressed in: EDPB and EDPS, 'EDPB-EDPS Joint Opinion 5/2021 on the proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)' (EDPB-EDPS, 18 June 2021).

in such a conduct.[121] The digital rights organization European Digital Rights has made a useful, long overview of the many missed opportunities to set a gold standard for AI use in the law enforcement area (and others), to which we have referred.[122] '*For those advocating for human rights, the AI Act is a mixed bag*' commented Laura Lazaro Cabrera, CDT Europe's Counsel and Director of the Equity and Data Programme. For Cabrera, the AI Act '*is only the beginning*' of the regulation needed.[123]

Overall, contrary to the US approach (in particular, with the AI 2024 Policy-measure)[124] that does *not* exempt, but imposes rigorous duties on law enforcement, the Act seems to be police-friendly, conditionally permitting critical AI-considerations that could have a hostile impact on fundamental rights and freedoms.[125]

In support of our argument concerning the AI Act's police-friendliness, two further comments can be made: one, on the AI Act's classification approach; and another, on transparency and accessibility (also in light of relevant provisions of the LED).

First comes the AI Act's classification-approach. While the Act makes distinctions on the basis of risk,[126] there is no guarantee that a given AI-use by law enforcement will actually qualify and be classified as a high-risk AI practice. For the purpose of the present discussion one can take examples of high-risk AI-implementations in the context of law enforcement that are aimed at: evaluating '*reliability of evidence in the course of the investigation or prosecution of criminal offences*'[127] (own emphasis); and '*profiling of natural persons as referred to in Article 3(4) of Directive (EU) 2016/680 in the*

---

[121] AI Act, 2024, art. 5, para. 1, lit. (d); Annex III, paras. 6 and 8. *See* also a discussion on risk assessments in: G van Dijck, *Predicting Recidivism Risk Meets AI Act*, 28 EUROPEAN JOURNAL ON CRIMINAL POLICY AND RESEARCH 407 (2022).

[122] EDRI, *EU's AI Act fails to set gold standard for human rights* (EDRI, 3 April 2024) *available at:* https://edri.org/our-work/eu-ai-act-fails-to-set-gold-standard-for-human-rights/ (last visited 25 Nov, 2024).

[123] Centre for Democracy & Technology Europe, *Landmark EU AI Act Sets Benchmark for AI Regulation, but Fails to Meet the Bar on Human Rights Protection*, (CDT Europe, 13 March 2024), *available at:* https://cdt.org/press/landmark-eu-ai-act-sets-benchmark-for-ai-regulation-but-fails-to-meet-the-bar-on-human-rights-protection/ (last visited 25 Nov 2024).

[124] Executive Office of the President, 'Memorandum for the Heads of Executive Departments and Agencies' (n 5).

[125] A discussion on the AI Act's vague and broad provisions that might fail to address criminal procedure's specifics in: R Stoykova, K Porter & T Beka, *The AI Act in a Law Enforcement Context: The Case of Automatic Speech Recognition for Transcribing Investigative Interviews,* SSRN (2024) *available at:* https://ssrn.com/abstract=4913090 (last visited 24 Oct, 2024).

[126] Assuming there are clear and reviewable criteria supporting the AI Act's risk classification. *See* above on the lack of clear reviewable criteria: *Supra* note 115.

[127] AI Act, 2024, Annex III, para 6, lit c.

course of the detection, investigation or prosecution of criminal offences'[128] (emphasis added).

The former category clearly covers the evaluation of evidence at the investigation stage but, at this early phase, one cannot tell whether such evidential materials will actually be used at trial. Moreover, it is unclear whether the 'reliability' evaluation covers AI-uses targeted at assessing the probative value of the evidence. And, in any event, many AI-practices might fall outside the scope of this (AI Act's) provision, given that, one might claim that they are deployed as mere data-searching/analysing instruments to aid investigative authorities.[129]

The latter category refers to AI-uses aimed at 'profiling'.[130] The question that naturally follows is whether profiling covers the mere sorting into a group on the basis of various criteria, like age or gender. The LED's definition of 'profiling' mentions examples, including past, present as well as future ('*predict*') conduct or characteristics and this seems to be limited by the AI Act's reference to '*profiling*' that occurs '*in the course of the detection, investigation or prosecution of criminal offences*'.[131] However, in the absence of a clear definition of 'detection' of criminal offences (at least, in some jurisdictions), it might be unclear whether and the degree to which AI-implementations aimed at, for instance, extracting patterns, which (patterns) may then be used without AI to identify criminal offences (*e.g.*, to map certain groups), can fall within the scope of the AI Act (Annex III, para 6, lit e).[132]

Second come transparency and accessibility, in view of both the LED and the AI Act. In general, access to information, in the light of fair trials, should cover access to materials at trial, but also evidential materials relevant for assessing reliability and lawfulness of expert evidence presented during trial, as well as information related to the way in which a conclusion was reached. In this regard, the LED (article 13) sets out a minimum content of information to be made available and it requires that competent authorities give further information to the data subjects in specific cases. And the AI Act (article 26, paragraph 11) sets out the duty of deployers of high-risk AI systems making decisions or assisting in decision-making to inform natural persons that they are subject to the use of that high-risk AI system, notably, for law enforcement-goals, the AI Act makes clear reference to the above article 13 of the LED. In addition, under article 86 of the AI Act, law enforcement entities, as deployers, must further give clear/meaningful explanations to the natural persons

---

[128] AI Act, 2024, Annex III, para 6, lit e.

[129] *Supra* note 13 at 213-215.

[130] *That is 'any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person, in particular to analyse or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behaviour, location or movements' (own emphasis). See*: Law Enforcement Directive,2016, art 3, point 4.

[131] AI Act, 2024, Annex III, para 6, lit e.

[132] *Id; See*, in more detail: *Supra* note 13.

after a decision concerning them was taken on the basis of the output of a high-risk AI system that has legal effects or significantly affects them.[133] While a minimalist interpretation of the AI Act and the LED's aforementioned provisions could mean that transparency-related obligations may be restricted to offering basic instructions on how to interpret the AI's output, perhaps a more expansive reading of these legal provisions should be adopted. Namely, law enforcement authorities should, not only ensure they know how to properly use the AI's output but also, explain how the system itself works. In other words, law enforcement authorities should always be in a position, as deployers of AI, to offer deeper insights into AI's functioning (by disclosing information on algorithms, but also details on the model's architecture and the data used for training).[134]

XII

The UN Toolbox for AI and the Collingridge Dilemma of Regulation

A 2021 Report on the risk of AI of the United Nations High Commissioner for Human Rights[135] investigates four key sectors in which the use of AI is a cause of concern due to its adverse impact on human rights. One of them is the criminal law area. Suggesting how to tackle the manifold challenges the AI poses, in the report, the High Commissioner emphasises that only a comprehensive human rights-based approach can provide a *toolbox* to find ways to minimise or avoid harm making the best out of technology for all. That toolbox contains four kinds of starting points (or *layers*), for regulation like proportionality and human rights due diligence.[136] The

[133] This can allow individuals to comprehend the role that the high-risk AI system may play in decision-making and the main elements of that decision. *See*: AI Act,2024, recital 171.

[134] Elsewhere, we have analysed in more detail this issue: P. De Hert, *al. el.*, , '*Evidence Law, Data Protection Standards and the Use of AI for Investigation and Prosecution*' in K Ligeti & N Visco AI EVIDENCE AND CRIMINAL PROCEEDINGS (Hart Publishing forthcoming).

[135] United Nations High Commissioner for Human Rights, *The Right to Privacy in the Digital Age: Report of the United Nations High Commissioner for Human Rights* (United Nations, 13 September 2021), *available at:* https://documents.un.org/doc/undoc/gen/g21/249/21/pdf/g2124921.pdf?OpenElement (last visited 25 Nov, 2024). *See*: C Cocito and P De Hert, *United Nations: AI can pose risks to human rights and privacy*, 174 PRIVACY LAWS & BUSINESS INTERNATIONAL REPORTS 13 (2021).

[136] Developments and deployments of AI should be surrounded by a framework of human rights principles, including equality, non-discrimination, participation, and accountability (layer 1). Principles of legality, legitimacy, necessity and proportionality shall infuse its application, particularly when AI is expected to restrict the right to privacy. Sometimes, like in the case of disproportionality or better alternatives, less interfering measures, or no AI, could be required (layer 2). AI applications should aim at ensuring availability, affordability, accessibility and quality as central elements for the realisation of economic, social and cultural rights (layer 3). There is a permanent requirement for human rights due diligence, a cyclic process through

Contd…

outcome of this would be the development of a system of laws, safeguards and oversight. In this system, data privacy laws remain the front-line framework to prevent uses of harmful AI, given that they are constantly adjusted and tightened, if necessary. Second category comprises of other laws and sector-specific legislation, in our case this would be about criminal procedure, evidence and defence rights. For the Commissioner, specific AI tools that cannot comply with human rights requirements may have to be prohibited under a legislative risk-based approach and moratoriums should be placed on potentially high-risk technologies before they can be assessed and have safeguards implemented. Furthermore, export control regimes should be strengthened to prevent cross-border trade of technologies that could violate human rights.

The Report suggests a multilayer oversight mechanism to be developed. Agencies should be forced to create meaningful human supervision and decision-making mandatory when AI is likely to have an adverse effect on human rights. Independent data privacy oversight bodies, aided by a cross-sectoral combination of administrative, judicial, quasi-judicial and/or parliamentary oversight bodies, and civil society engagement are paramount to challenge the complexities, opacity and power asymmetries of the global data environment.

Recently, Robroek defended a step-by-step approach towards regulating AI.[137] According to Robroek, AI is not just 'a' technology, but a game changer (a 'genuine new system technology') that confronts us with new uncertainties and the famous Collingridge dilemma about initially knowing not enough (information deficit) and later knowing too much, but too late (power deficit).[138] Robroek wants to go beyond the dilemma, by regulating at an early phage but not *via* legislation, but guidance developed by the public prosecutor office on how to handle AI in court that is subsequently reviewed in court. 'Hard' legislation can follow later.

To a certain degree, having in mind possible abuses,[139] Robroek's model of regulation can be applied. A good example is the issue of reliability of AI. We have

---

which human rights risks are identified, assessed and prevented, when AI systems are acquired, developed, deployed and operated, as well as before individuals' data are shared or used. If due diligence establishes that AI is incompatible with human rights law and no meaningful means exist to mitigate harm, it should be banned (layer 4). Due diligence also requires states to engage in meaningful consultations with AI-affected stakeholders, primarily vulnerable groups.

[137] P Caeiro, M Caianiello, J Spencer and K Ambos (eds), *Elgar Encyclopedia of Crime and Criminal Justice* (Edward Elgar Publishing, 2024).

[138] Collingridge's basic insight was that we can regulate a given technology when it's still young and unpopular, but there is not always enough understanding and there might be unanticipated and undesirable consequences (information deficit); or we can wait and see what those consequences are, but then risk losing control over its regulation, because practices have settled often to the advantage of the powerful who do not want to give in on acquired advantages (power deficit).

[139] Soft law and self-regulation cannot be used by law enforcement authorities to grant themselves

Contd…

discussed *above* several problems posed by AI in terms of reliability: authenticity and integrity of the 'raw' digital data; or hidden miscodes of opaque AI systems and humans operating these tools jeopardising the reliability of AI evidence. Quezada-Tavarez *et. al.* who have extensively worked in this area, see a first solution in the chronological documentation of evidence and full respect of existing soft law's guidance, such as the 2013 Council of Europe's Guide on Electronic Evidence and other supervisiory documents.[140] Helga Molbæk-Steensig and Alexandre Quemy discuss the 2018 European Ethical Charter on the use of Artificial Intelligence in Judicial Systems and their Environment and first pilot projects set up by the Council of Europe.[141]

Additional guidance at domestic level *via*, for instance, instructions of the Public Prosecutor's Office as Robroek suggests, could be helpful granted that they respect the European guidance.

XIII

Ligue des droits humains (2022): Rejecting Machine-Learning Algorithms

We saw *above* that, contrary to the US, Europe seems to lack a harmonised (federal-like) system with specific rules (*e.g.*, governing exclusion of illegally obtained

---

new investigatory powers. This is precisely what the Dutch authorities are doing today with facial recognition. In early 2024, the press reported that the minister had given the police permission to experiment with facial recognition software. An internal police review committee assesses each application to see if this can be allowed. The whole was set up without any consultation of the Dutch Personal Data Authority. The argument of police forces, public prosecutors and the political responsible Minister is that courts will have an opportunity to speak out on the evidence in court. But that kite doesn't fly, argues Professor of Law and Data Science, B Custers. '*The problem is that in many cases* (…) *don't end up in court. Especially precisely in a situation where someone is not guilty or has done nothing wrong, it never comes to a case in court. But in the meantime, people's rights have been violated*', M Schrader, Y Vugts & A Mae, *Police experiments with facial recognition, but legislation lacking,* Politie experimenteert met gezichtsherkenning, maar wetgeving ontbreekt, 5 January 2024, *available at:* https://nos.nl/nieuwsuur/artikel/2503831-politie-experimenteert-met-gezichtsherkenning-maar-wetgeving-ontbreekt (last visited 25 Nov, 2024).

[140] Council of Europe, 'Electronic Evidence Guide. A Basic Guide for Police Officers, Prosecutors and Judges' (Council of Europe 2013); European Union Agency for Cybersecurity (ENISA), 'Electronic Evidence - a Basic Guide for First Responders' (2015). These documents contain requirements for a solid chain of custody both at moment of collection of 'raw' digital data and at the moment of subsequent processing of these data by AI tools. *See*: *Supra* note 34 at 541.

[141] *See* the CoE's five general principles in the 2018 European Ethical Charter on the use of Artificial Intelligence in Judicial Systems and their Environment. *See* for a discussion, Helga Molbæk-Steensig & Alexandre Quemy, *Artificial Intelligence and Fair Trial Rights* in ARTIFICIAL INTELLIGENCE AND HUMAN RIGHTS, 265-280 (Alberto Quintavalla & Jeroen Temperman, ed. 2023).

evidence)[142] to be followed within its national, often divergent procedures. In addition to the absence of a robust, clearcut and harmonised mechanism, Europe also lacks *case law*s on AI-related evidential materials. Both the ECtHR and the CJEU are pouring out strings of judgement that define and identify safeguards regarding surveillance, data retention and other practices. Of course, there is rich case law on due process that is generic and applicable to AI-related work. But there is no relevant ruling by the ECtHR specifically on the use of AI based evidence in courts.[143]

On the CJEU side, there is *Ligue des droits humains* (2022) concerning Directive 2016/681 on passenger name record data (PNR Directive).[144] Although this case concerns the use of AI by law enforcement to detect 'suspicious passengers', the CJEU identifies a set of criteria to be met when using AI, that can help courts to assess admissibility of evidence obtained. The CJEU requires that the system can only be applied to offences that have -at least indirectly- an objective connection with the purpose of protecting air transport of passengers and not for purposes other than those explicitly mentioned in the law. The CJEU then goes on to

---

[142] *See*, for instance, a discussion on article 32 of the Preliminary Title of the Belgian Code of Criminal Procedure, where the authors stress that limiting the exclusionary rule via an exhaustive list of three grounds might be unreasonable, in: F Verbruggen & C Conings, *After Zigzagging Between Extremes, Finally Common Sense? Will Belgium Return to Reasonable Rules on Illegally Obtained Evidence?*, 7(1) IBRASSP (2021), *available at:* https://www.redalyc.org/journal/6739/673972096009/html/ (last visited 24 Oct, 2024). *See* also: Fair Trials, *Unlawful evidence in Europe's courts: principles, practice and remedies* (Fair Trials, October 2021), *available at:* https://www.fairtrials.org/app/uploads/2021/11/DREP-report.pdf (last visited 24 Oct, 2024).

[143] The only directly related case so far is the *Glukhin* v. *Russia,* (2022), App. No. 11519/20. Judgment (2023) rejection of Russian use of facial recognition that was insufficiently regulated in terms of limitations, categories of crimes, use of sensitive data and guarantees. *See,* in more detail: M Zalnieriute, *Glukhin v. Russia. App. No. 11519/20. Judgment*, 117(4) AMERICAN JOURNAL OF INTERNATIONAL LAW 695, 701, (2023). As the author concludes, this judgement, though the first on face recognition mass-surveillance, reinforces the Court's long-adopted liberal view that the state may keep on using mass surveillance, where certain procedural guarantees are in place. It accepts that such surveillance is legitimate to use in public spaces; albeit, the judgement is also emblematic of the ECtHR's more restrictive approach to more authoritarian states, like Russia, and the difficulty of the ECtHR in reinforcing a permissive approach to mass surveillance in the face of growing evidence that the procedural guarantees are insufficient to avoid human rights violations. *See* also: C Cocito, *Glukhin v. Russia: Facial Recognition Considered Highly Intrusive but Not Inconsistent with Fundamental Rights,* (Strasbourg Observers, 9 January 2024), *available at:* https://strasbourgobservers.com/2024/01/09/glukhin-v-russia-facial-recognition-is-a-highly-intrusive-technology-but-the-court-abstains-from-considering-a-potential-inconsistency-with-fundamental-rights/ (last visited Feb 25, 2025). Finding that the ECtHR's elusive and vague approach '*leaves a number of unanswered concerns, one of which is whether such technologies are intrinsically compliant with human rights* (…) *With this ruling, the Court may have missed another salient chance to position itself more firmly within the debate concerning mass-surveillance technology's compliance with fundamental rights*'.

[144] *Ligue des droits humains* v. *Conseil des ministres* (Judgment of the Court, Grand Chamber, 21 June 2022), CJEU, Case C-817/19, .

formulate important requirements on retention periods,[145] and most relevant for us, the requirements on due process, transparency, and the quality of the data.[146]

Surprisingly, little has been written on *Ligue des droits humains*. In professional discussions, we have noted, that the CJEU's rejection of the use of machine-learning algorithms did not please all those involved in writing about AI. It is clearly not ready for hard regulation yet. In our view, it is here that the proposal for a moratorium defended by the United Nations High Commissioner for Human Rights, in its 2021 report on AI (discussed *above*), makes most important point. This idea is one of placing a stop on potentially high-risk technologies (not all) before they can be assessed and have safeguards implemented. The idea makes sense considering the research of Brandon Garrett, who finds little reasons for black box AI systems in the criminal law-sphere and holds that interpretable AI systems perform as well and have clear advantages in terms of due process rights (*above*). In any event, whether and how the CJEU's views will be followed remains unclear, in light of AI-related expectations, its increased use and its understandability/controllability-limitations (as analysed above).[147]

With this caveat, we think there is no reason to invoke an information deficit as an argument not to intervene during justice procedures. The CJEU's insistence on higher legal (due process) protection, transparency and data quality requirements make eminent case for immediate regulation. The regulation, in the first phase may be initiated, either, through binding norms made by the criminal authorities

---

[145] Data should in principle not be retained for longer than six months, unless it concerns data of persons in respect of whom there are objective indications that there is a risk that they will commit terrorist offences or serious crimes at least in some connection with the flight.

[146] *Firstly*, to strengthen due process and transparency, the Court requires that high standards of legal protection must be met and that agencies ensure that individuals are always able to learn about the reasons and evidence based on which a decision was made, including the predetermined assessment criteria and the operation of the programs that use these criteria. *Secondly*, to ensure high quality of the data used to profile and to avoid risks of discrimination, the five following requirements are to be respected: 1) data should undergo a prior assessment; 2) demanding requirements are placed on the automated processing. Machine-learning algorithms cannot be used, since these can adjust the assessment process without human intervention or control. Those adjustments mainly concern the assessment criteria on which the results of that process are based and the weighing of those criteria. 3) pre-determined criteria should be chosen to specifically target individuals, who may be reasonably suspected of involvement in terrorism or serious crimes, to consider both incriminating and exculpatory elements and not to create direct or indirect discrimination. This means that predetermined criteria must be chosen in such a way that, even if worded neutrally, they cannot in practice be particularly disadvantageous to persons with the protected characteristics; 4) databases with which the PNR information is compared must themselves be non-discriminatory; 5) throughout the processing, high standards must be applied to rule out any discriminatory results, especially given the relatively large number of false positives that can be generated.

[147] *See* above sections on case law of the CJEU and the ECtHR, with further references to critical points made by J Van Banning, THE AI ACT AS A SAFEGUARD FOR EQUALITY OF ARMS IN DATA-DRIVEN CRIMINAL INVESTIGATIONS, 212(2025).

themselves reviewed in court or immediately *via* hard law at national or European level. Such a regulation, may be based on the principles set out by the UN High Commissioner, e.g., cyclic human rights diligence (*above*), in particular the principle, that *the higher the risk for human rights, the stricter the legal requirements the use of AI technology must satisfy* (*above*).

XIV

Improving European Preparedness for AI in the Criminal Area

Compared to the US legal system that we studied in a previous paper, the European legal situation seems to offer a similar protection, regarding evidence and due process, at least, on the basics parameters. Europe appears less harmonised, with national divergencies leaving no room for a clear-cut and robust (federal-like) scheme. On equal protection and non-discrimination, the European framework could be seen as more defence-friendly, absolving persons, who have been discriminated against, from showing discriminatory intent. Yet, Europe has no concrete, standardised scrutiny levels applicable when ascertaining undue discrimination. Regarding the legality principle and the presence of hard law on AI transparency and privacy, the US scheme seems more targeted at technological usages that are subjected to brightline rules letting citizens know and foresee what is permitted and what is prohibited in specific contexts. The US appear less tolerant toward law enforcement that (contrary to Europe) are expressly *not* exempted from stringent requirements and obligations.

Furthermore, our analysis of case law of the European Courts illustrate that critical questions remain open (*e.g.*, on the qualification of certain AI-uses as expert evidence, their possible opaque and secret use or (in)applicable reliability-/contestability-related standards). The same is true for the AI Act, whose interplay with the LED can reveal many missed opportunities.

In light of the foregoing considerations, Europe's preparedness and readiness could be improved. Europe should aim to reach the US high-levelled concreteness: that is, do-able if policy- and decision-makers relied upon, dynamically interpreted and effectively applied the European legality principle and the principle of legal certainty with the objective of engaging in brightline rulemaking, addressing and tackling specific data- and AI-related concerns. This is no less than the core message of *Ligue des droits humains*. It is also voiced in literature by Quezada-Tavarez *et al.* who call for policy-making and new laws at the European and domestic level, for transparent architecture of law enforcement regulating AI and for training of professionals. More effective protection of citizens against potential abuses of AI

would be in reach, if Europe adopted a (US-like)[148] critical, less tolerant and less permissive approach to technological uses by law enforcement or other state authorities.[149]

The moment seems to be ripe for regulatory improvements at the European level and in the respective domestic laws. The University of Luxembourg, led by Professor Ligeti, has presented a first comprehensive study on the use of AI in court.[150] The research identifies three areas of possible improvement of the legal framework with regard to the use of AI in criminal procedure: strengthening reliability requirements (for instance requiring that the use of AI evidence needs additional evidence); strengthening human oversight; and strengthening defence rights including the right to an expert. The minds seem to be ripe too. Ligeti identifies a recent tendency (very similar to what we found in our previous work on the US) to move from blind trust to real scrutiny by the courts. That is a positive sign.

Part of the program to strengthen the existing framework is already activated in court (the CJEU's *Ligue des droits humains*; the ECtHR's *acquis* on fair trial and insistence on the adversarial principle concerning expert evidence in *Quadrature*). Another source of activation comes from the AI Act (influenced by data protection). Further, strengthening regulations should give due consideration to five novel defence rights identified by the Luxembourg research project (Ligeti); the proposed starting points for the use of expert testimony (Gless), together with other proposals for novel or updated rights by Stoykova and Palmiotto.

Only recognition of novel rights and codification of novel interpretations of existing rights can address transformations of the traditional procedure, caused by AI, and preserve the core principles of a fair trial.

A key message of this contribution is that Europe can and should rely upon the legality principle and the principle of legal certainty with a view to establishing (most probably, at national level) brightline rules, capable of tackling concrete data-

---

[148] Our insistence on certain aspects of the US approach to AI regulation has to be comprehended under the light of some obstacles to AI governance (as detected in the US), such as the political challenge to adopt a comprehensive federal law offering a consistent regime for AI governance; the enhanced protection of freedom of expression (that can, in the US, cover various AI/data-related operations, from gathering and disseminating publicly available data to generating new content); or the preference of innovation over precaution (that may, in some instances, absolve firms from liability concerning user-generated content). On these and similar obstacles to comprehensive AI (privacy) governance in the US, *see*: Elana Zeide, *Expanding the Paradigm: Generative Artificial Intelligence and U.S. Privacy Norms*, 1 CAMBRIDGE FORUM ON AI: LAW AND GOVERNANCE 9-10 (2025).

[149] While the US may traditionally be regarded as a paradigmatic example of police-friendliness, in their 2024 AI Policy 'Memorandum for the Heads of Executive Departments and Agencies' (n 5), they seem to impose rigorous duties on law enforcement.

[150] *Supra* note 31.

/AI-related concerns in this specific area. Whilst moving toward such brightline rulemaking, regulators could very well learn from recent initiatives around the globe, aimed at making the best out of AI, bringing about efficiency and effectiveness in and, thus, transforming the criminal justice system; but also tackling AI's limitations (*e.g.*, errors and bias) that may undermine fairness and human rights and freedoms.

By way of example, we refer to Colombia's AI Guidelines for the Responsible and Safe Use of Generative AI in the Judicial Branch, targeted at upholding fundamental rights and judicial integrity.[151] Colombia's AI Guidelines were launched in December 2024, upon cooperation between UNESCO and Colombia's Superior Council of the Judiciary. Within its 'AI and the Rule of Law programme',[152] UNESCO had published its guidelines for the use of AI systems in courts, with a view to providing concrete guidance for organisations and individuals forming part of the judiciary; and the ultimate objective of guaranteeing compliance of AI-implementations with core fundamental principles of justice, human rights and the rule of law.[153]

Relying and building upon UNESCO's guidelines, Colombia seems to go beyond the establishing and promotion, in theory, of known *and* novel ethical principles (from transparency and explainability to non-substitution of human rationality and attribution)[154] by providing for *practical* guidance on how to ethically adopt AI in

---

[151] COLOMBIA'S GAZETTE OF THE JUDICIARY AND ITS SISTEMA DE INFORMACIÓN DE RELATORÍA PRESIDENCIA, *Colombia's AI Guidelines*, *available at:* https://actosadministrativos.ramajudicial.gov.co/web/Acto%20Administrativo/Default.aspx?ID=19280 (last visited 8 April 2025).

[152] UNESCO, *AI and the Rule of Law: Capacity Building for Judicial Systems*, *available at:* https://www.unesco.org/en/artificial-intelligence/rule-law/mooc-judges (last visited 6 May 2025); *Artificial intelligence and the rule of law*, *available at:* https://unesdoc.unesco.org/ark:/48223/pf0000390708#:~:text=Why%20AI%20and%20the%20Rule,human%20rights%20and%20ethical%20concerns. (last visited 6 May 2025).

[153] UNESCO guidelines established thirteen principles for the use of AI in the judiciary: protection of human rights; proportionality; safety; information security; awareness and informed use; transparent use; accountability and auditability; explainability; accuracy and reliability; human oversight; human-centric design; responsibility; multi-stakeholder governance and collaboration. These principles had been further elaborated in: UNESCO's Recommendation on the Ethics of Artificial Intelligence, *available at*: https://unesdoc.unesco.org/ark:/48223/pf0000381137 (last visited 6 May 2025). *See* also: UNESCO, *Global Toolkit on AI and the Rule of Law for the Judiciary*, *available at:* https://unesdoc.unesco.org/ark:/48223/pf0000390708#:~:text=Why%20AI%20and%20the%20Rule,human%20rights%20and%20ethical%20concerns. (last visited 6 May 2025).

[154] *See* Colombia's AI Guidelines, 2024, art 3, referring to fifteen overarching principles and guarantees: primacy of fundamental rights; ethical regulation; adaptation to best practices and collective standards; monitoring, improvement and continuous adaptation; collaborative and multi-stakeholder governance; non-substitution of human rationality and attribution; adequacy; responsibility and informed use; supervision, control and human verification; risk prevention;

Contd…

various cases (*e.g.*, they categorise applications based on risk levels, from routine administrative tasks to more sensitive applications involving case law),[155] but also by imposing concrete duties on stakeholders in he judiciary (*e.g.*, to clearly disclose whether, how and what AI-tools are deployed).[156] Notably, Colombia moved toward such ethical AI-regulation, after a Constitutional Court ruling on risks stemming from unregulated use of generative AI by judges (*e.g.*, copying/pasting ChatGPT's responses to case-related prompts)[157] that might interfere with the right to a fair trial. While AI cannot replace human judgement (judicial reasoning), its potential to address specific societal needs for justice (like the large number of cases pending before courts) should not be ignored.[158]Colombia's initiative and the work of UNESCO that paved the way could work as a baseline and point of reference for adequate AI-regulation in Europe. The legal regims on AI, in the latter, as our analysis revealed, appears to miss, even in theory, a Colombian-like establishment of regulations/mandatory principles going beyond mere ethical principles[159] lack stringent obligations including outrightly ban in certain cases.[160] In our view, no less than a profound *detox* is needed to safeguard, both *in theory and in practice*, the people's interest in becoming aware of the fact that the state uses AI, which is a technology, that can be surrounded by unfathomable black boxes and does produce output that unfairly results into discriminative decisions and content and that may not always deliver humanly intelligible explanations.

---

equality and equity; personal data protection and privacy; information security; explainability; and transparency.

[155] *See* Colombia's AI Guidelines, 2024, art 4 (4.1 to 4.3), distinguishing between: AI-uses in the administrative/judicial management (e.g., email-drafting); AI-implementations requiring a more detailed review (e.g., compiling data related to a case under study); and AI-applications demanding enhanced transparency (like simulation of case-specific decision scenarios).

[156] *See*, for instance, Colombia's AI Guidelines, 2024, art 8, enlisting various obligations of officials and employees of the judicial branch, who use AI tools for the performance of their duties.

[157] Juan David Gutiérrez, *ChatGPT in Colombian Courts: Why we need to have a conversation about the digital literacy of the judiciary*, (VERFBLOG, 23 February 2023), *available at:* https://verfassungsblog.de/colombian-chatgpt/ (last visited 6 May, 2025).

[158] *See*: UNESCO, *Justice meets innovation: Colombia's groundbreaking AI guidelines for courts*, (UNESCO, 25 February 2025; last update 1 April 2025), *available at :* https://www.unesco.org/en/articles/justice-meets-innovation-colombias-groundbreaking-ai-guidelines-courts#:~:text=Colombia%20has%20achieved%20a%20groundbreaking,safeguarding%20ethics%20and%20human%20rights. (last visited 8 April, 2025).

[159] It is reminded that, concerning the AI Act, the European Parliament's proposal for an article (4a) on general principles applicable to all AI systems was abandoned. *See* Amendments Adopted by the European Parliament on 14 June 2023 (COM (2021)0206 – C9-0146/2021 – 2021/0106(COD)).

[160] It is reiterated that, with regard to the AI Act, the lack of clear bans was stressed by the EDPB and the EDPS in: 'EDPB-EDPS Joint Opinion 5/2021 on the proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)' (EDPB-EDPS, 18 June 2021).