Himachal Pradesh National Law University, Shimla (India)

SLR
Shimla Law Review

A UGC CARE Listed Journal

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

Volume: VI (2023)

SPECIFIC LAWS GOVERNING USE OF AI IN CRIMINAL PROCEDURE AND ATTENTIVE CRIMINAL JUDGES: American Songbook for Global Listeners

Paul De Hert & Georgios Bouchagiar

https://doi.org/10.70556/hpnlu-slr-v6-I1-2023-01

This article can be downloaded from: https://www.hpnlu.ac.in/journal-level-3.aspx?ref-id=18

Recommended Citation:
Paul De Hert & Georgios Bouchagiar *Specific Laws Governing Use of AI in Criminal Procedure and Attentive Criminal Judges: American Songbook for Global Listeners,* VI SML. L. REV. 1 (2023). https://doi.org/10.70556/hpnlu-slr-v6-I1-2023-01

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 VI *ISSN: 2582-1903* *April 2023 - March 2024*

Special Article *Page*

Articles

Notes and Comments

SPECIFIC LAWS GOVERNING USE OF AI IN CRIMINAL PROCEDURE AND ATTENTIVE CRIMINAL JUDGES: American Songbook for Global Listeners

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

[Abstract: *Artificial Intelligence (AI) can be used in the criminal justice system to support human decision-making at various stages of the proceedings. Despite heavy criticism on AI's opacity, complexity, non-contestability or unfair discrimination, such AI-implementations are often favoured, in light of alleged accuracy, effectiveness or efficiency in the overall decision-making process. After briefly recalling some key functions of AI in criminal procedures, the paper addresses whether and the degree to which AI-uses can comply with the United States (US) Federal Rules of Evidence and the constitutional rights to due process, equal protection and privacy. Recent case law (e.g., Puloka and Arteaga) and legal initiatives, such as the 2024 AI Policy and California's 2024 Rules of Court, are also discussed. This paper ends with five important take-homes for global readers and regulators intending to introduce AI into their jurisdictions.*]

Keywords: *Artificial Intelligence, law enforcement, criminal justice, evidence, due process, equal protection, and privacy etc.*

I

AI-Made Decisions and Contents: Promises and Limitations

AI can analyse large amounts of data in a fast and sophisticated way and deliver an output (*e.g.*, a decision, a likelihood or content). As such, AI is said to be particularly attractive to law enforcement and criminal justice authorities that can be assisted in,

[*] *Paul De Hert* is a full professor and Vice-Dean of the Faculty of Law & Criminology at the Vrije Universiteit Brussel (VUB) and director of the Brussels Fundamental Right Research Centre (FRC). He is also associated professor at Tilburg University. De Hert is Member to the Scientific Committee of the European Union Agency for Fundamental Rights (FRA) based in Vienna. *Email*: paul.de.hert@tilburguniversity.edu

[**] Georgios Bouchagiar is a postdoctoral researcher at the Vrije Universiteit Brussel (VUB). *Email*: georgios.bouchagiar@vub.be

among others, investigating crime, charging or convicting suspects[1] or evaluating defendants at various stages of the criminal proceedings (*e.g.*, at the pretrial,[2] the sentencing[3] or the post-sentencing[4] phase).

To its advocates, AI could increase statistical precision[5] and enhance efficiency, effectiveness and accuracy of decision-making processes, as well as reduce workload of state authorities. AI's features, like advanced analytics, could in fact be used to investigate and prevent offences and guarantee enhanced public safety and increased confidence in law enforcement and the criminal justice system, in general.[6]

[1] *See*, Cybercheck in: T Stelloh, *An AI tool used in thousands of criminal cases is facing legal challenges*, NBC NEWS (May 3, 2024) *available at*: https://www-nbcnews-com.cdn.ampproject.org/c/s/www.nbcnews.com/news/amp/rcna149607 (last visited Sep. 8, 2024); The Law Reporters, *Legal Challenges Mount Against AI Software Used in Thousands of Criminal Cases* (7 May 2024) *available at*: https://thelawreporters.com/legal-challenges-mount-against-ai-software-used-in-thousands-of-criminal-cases/ (last visited Sep. 8, 2024).

[2] *See*, A Novokmet, Zv Tomičić & Z Vinković, *Pretrial Risk Assessment Instruments in the US Criminal Justice System — What Lessons Can Be Learned for the European Union* 30(1) INTERNATIONAL JOURNAL OF LAW AND INFORMATION TECHNOLOGY 1 (2022).

[3] *See generally*, J Ryberg & J V Roberts (eds.), SENTENCING AND ARTIFICIAL INTELLIGENCE 122 (Oxford University Press 2022).

[4] A popular tool, used at various stages of the criminal proceedings, is COMPAS. An analysis in: C Rudin, C Wang & B Coker, *The Age of Secrecy and Unfairness in Recidivism Prediction* 2(1) HARVARD DATA SCIENCE REVIEW (2020); E Jackson & Chr Mendoza, *Setting the Record Straight: What the COMPAS Core Risk and Need Assessment Is and Is Not* 2(1) HARVARD DATA SCIENCE REVIEW (2020).

[5] From a mathematical point of view, AI, capable of processing large amounts of data, could make more correlations that could (not by necessity, but still could) increase statistical precision. On precision and recall, see in more detail: ICO, *What do we need to know about accuracy and statistical accuracy?* (Mar. 15, 2023) *available at*: https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/artificial-intelligence/guidance-on-ai-and-data-protection/what-do-we-need-to-know-about-accuracy-and-statistical-accuracy/ (last visited Sep. 8, 2024); ChatGPT Guide, *What is Precision: Artificial Intelligence Explained* (Feb. 26, 2024) *available at*: https://www.chatgptguide.ai/2024/02/26/what-is-precision-artificial-intelligence-explained/#:~:text=Understanding%20Precision%20in%20AI,model%2C%20particularly%20in%20classification%20tasks. (last visited Sep. 8, 2024); S Alaoui, *Statistical Evaluation: Unveiling AI's Performance and Precision*, MEDIUM (Sep. 20, 2023) *available at*: https://medium.com/aimonks/statistical-evaluation-unveiling-ais-performance-and-precision-31bc5f0d5600 (last visited Sep. 8, 2024).

[6] *See*, Chr Rigano, *Using Artificial Intelligence to Address Criminal Justice Needs* 280 NATIONAL INSTITUTE OF JUSTICE JOURNAL (2019).

However, AI can pose at least three critical challenges. *First,* AI's functioning seems to resist transparency and intelligibility:[7] AI tools can be subject to Intellectual Property (IP) rights, limiting or denying access to their autonomous (often non-human-guided) modus operandi; and, even where access is permissible, such modus operandi may be humanly incomprehensible.[8] *Second*, AI-implementations could have an unfairly discriminative impact; *e.g.,* constitutional concerns may be raised by the use of risk assessment technologies for sentencing goals that can lead to sorting defendants into groups on the basis of gender or other protected grounds.[9] *Third*, there is a practical concern that should deeply preoccupy the entire criminal justice system: AI may suffer error rates, due to, among others, bias or inadequate review, training and validation.[10] These three risks can be –and have in some cases been– materialised, where AI is used to augment evidential materials (like a video) by adding content (footage),[11] to label a defendant as highly risky to reoffend on the basis of invalidated tools[12] or, in general, to deliver decisions and content, whose reliability and accuracy may be hardly scrutinised.[13]

[7] On complexity of AI decision-making, *see generally*, A Rubel, Cl Castro & A Pham, ALGORITHMS AND AUTONOMY: THE ETHICS OF AUTOMATED DECISION SYSTEMS (Cambridge University Press 2021); W Barfield, THE CAMBRIDGE HANDBOOK OF THE LAW OF ALGORITHMS (Cambridge University Press 2020).

[8] *See*, A Nishi, *Privatizing Sentencing: A Delegation Framework for Recidivism Risk Assessment* 119 COLUMBIA LAW REVIEW 1671 (2019).

[9] *See*, S Starr, *Evidence-Based Sentencing and the Scientific Rationalization of Discrimination* 66 STANFORD LAW REVIEW 803 (2014).

[10] While inclusiveness and training seem essential to eliminate/minimise bias, there may be other ways to tackle discrimination-related issues. See, for instance, discussions on the 'debiasing paradox' in: A R Martínez, *The Debiasing Paradox: What If Algorithms Do Not Deviate from Human Nature*, THE DIGITAL CONSTITUTIONALIST (Mar. 22, 2023) *availale at*: https://digi-con.org/the-debiasing-paradox-what-if-algorithms-do-not-deviate-fromhuman-nature/ (last visited Sep. 8, 2024).

[11] *See*, *State* v. *Puloka* (No 21-1-04851-2KNT) ('*Puloka*'), para 13 ('(…) The video evidence produced by the Topaz Video AI enhancement model does not satisfy ER 401, as the resulting video does not show with integrity what actually happened but uses opaque methods to represent what the AI model 'thinks' should be shown (…)').

[12] In fact, at least one court considered a risk assessment technology that had not been validated for the state's own population. See: *State* v. *Loomis*, 371 Wis 2d 235 (Wis 2016) ('*Loomis*') para 261.

[13] A critical discussion on ChatGPT in: Michael Townsen Hicks, James Humphries & Joe Slater, *ChatGPT Is Bullshit* 26(38) ETHICS AND INFORMATION TECHNOLOGY (2024) *available at*: https://doi.org/10.1007/s10676-024-09775-5 (last visited Sep. 8, 2024). Failure to scrutinise reliability and accuracy of data driven operations is an old subject of legal discussions; for a critique on big data policing, see: Andrew Guthrie Ferguson, THE RISE OF BIG DATA POLICING: SURVEILLANCE, RACE, AND THE FUTURE OF LAW ENFORCEMENT (NYU Press 2017) 177ff ('(…) [t]he promise of big data policing depends on the size and scale of the data

Contd…

Despite the above critique, the use of AI is embraced by law enforcement and criminal justice authorities tempted by the technology's promises. Given recent trends in the US, it is worth assessing whether and the degree to which AI-implementations can respect fundamental principles that must be in place at various stages of the criminal procedure.

In what follows, we first examine the US criminal procedure and the constitutional framework, revealing the way and the extent to which fundamental rights have been incorporated at the US state level (section II). Thereafter, we discuss the right to due process, requiring enhanced transparency of AI, so that the defence can effectively challenge allegations against it and their reliability (section III). The discussion moves naturally to the basic requirements of evidence law under the US Federal Rules of Evidence (section IV), as well as more explicit demands of expert evidence (section V), under the Federal Rule of Evidence 702 (as amended in 2023; section VI). Moreover, we address the equal protection clause, demanding strict-scrutiny of AI-uses that could interfere with fundamental rights (section VII). Then, we analyse the right to privacy, calling for brightline rulemaking targeted at specific AI-implementations to ensure, among others, integrity and reliability of personal data that may be involved in and affect the quest for truth (section VIII).

Overall, as our analysis of recent case law (*e.g.*, *Puloka* and *Arteaga*) and regulatory interventions (in particular, the 2023 amendments to the Federal Rule of Evidence 702) demonstrates, there has been a shift from obliging acceptance ('science bias' where judges used to accept everything) toward a more critical judicial scrutiny of AI-related evidence. In our conclusion (section IX), we summarise our findings and rely upon recent legal initiatives (especially, the 2024 AI Policy and California's 2024 Rules of Court) to offer readers take-homes based on the US analysis.

analyzed (…) [b]ut data holes remain because of systemic pressures on what type of data gets collected (…) [a]s society moves toward a more data-dependent policing system ,filling these data-holes or, at a minimum, acknowledging their existence can counteract a blind reliance on numerical, probabilistic suspicion (…)'). See also recent cases of using ChatGPT in court and referencing fake cases in Molly Bohannon, *Lawyer Used ChatGPT in Court—And Cited Fake Cases A Judge Is Considering Sanctions*, FORBES (Jun. 8, 2023) *available at*: https://www.forbes.com/sites/mollybohannon/2023/06/08/lawyer-used-chatgpt-in-court-and-cited-fake-cases-a-judge-is-considering-sanctions/ (last visited Sep. 8, 2024); Erroneous cell-phone tracking data brought as evidence in Martin Selsoe Sorensen, *Flaws in Cellphone Evidence Prompt Review of 10,000 Verdicts in Denmark*, THE NEW YORK TIMES (Aug. 20, 2019) *available at*: https://www.nytimes.com/2019/08/20/world/europe/denmark-cellphone-data-courts.html (last visited Sep. 8, 2024); Convictions after considering faulty software in Sachin Ravikumar, *What is Britain's Post Office scandal?* REUTERS (Jan. 11, 2024) *available at*: https://www.reuters.com/world/uk/what-is-britains-post-office-scandal-2024-01-09/#:~:text=WHAT%20IS%20THE%20POST%20OFFICE,showed%20money%20missing%20from%20accounts. (last visited Sep. 8, 2024).

II

Criminal Procedure and the Constitutional Framework[14]

The original American constitution of 1787 contains very few human rights; and the few that do appear are mainly concerned with criminal procedure.[15] In section 9, article I, the *Habeas Corpus* law is mentioned (the right to judicial approval for deprivation of freedom) and the ban on retrospectivity of laws. Article 3 contains guarantees in connection with the autonomy of the judicial administration and acknowledges the right to *trial by jury*.[16]

This list of human rights was considerably extended in 1791 with the incorporation of the first ten amendments to the constitution. These amendments form the well-known *Bill of Rights* and contain those specific human rights that presently appeal the most to the imagination, including several human rights to be used in criminal procedure.[17] The fourth amendment protects against unreasonable searching of premises by the police and in certain situations imposes the obligation of a warrant.[18] Similar importance is given to the fifth amendment, where rights are recorded, such as the right to judgement of a case by a grand jury, the right not to have to give evidence against oneself, the right not to be charged twice for the same crime and also the so-called '*due process of law*' right: '*No person shall be deprived of life, liberty, or property, without due process of law*'.[19]

[14] P De Hert, *Legal Procedures at the International Criminal Court. A Comparative Law Analysis of Procedural Basic Rights* in SUPRANATIONAL CRIMINAL LAW: A SYSTEM SUI GENERIS 79 (R Haveman, O Kavran and J Nicholls, eds., Intersentia 2003).

[15] T M Fielding Fryling, CONSTITUTIONAL LAW IN CRIMINAL JUSTICE 552 (Aspen Publishing, 2023).

[16] Article 3, section 2, clause 3: '*The Trial of all Crimes, except in Cases of Impeachment, shall be by Jury; and such Trial shall be held in the State where the said Crimes shall have been committed; but when not committed within any State, the Trial shall be at such Place or Places as the Congress may by Law have directed*'.

[17] A R Amar, AMERICA'S CONSTITUTION: A BIOGRAPHY 635 (Random House, 2006).

[18] See Fourth Amendment: '*The right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures, shall not be violated, and no Warrants shall issue, but upon probable cause, supported by Oath or affirmation, and particularly describing the place to be searched, and the persons or things to be seized*'.

[19] See Fifth Amendment: '*No person shall be held to answer for a capital, or otherwise infamous crime, unless on a presentment or indictment of a Grand Jury, except in cases arising in the land or naval forces, or in the Militia, when in actual service in time of War or public danger; nor shall any person be subject for the same offence to be twice put in jeopardy of life or limb; nor shall be compelled in any criminal case to be a witness against himself, nor be deprived of life, liberty, or property, without due process of law; nor shall private property be taken for public use without*

Contd…

For a long time, it was thought that the American constitution, including the amendments, were only applicable to federal authorities. Only these authorities, and not state authorities, were bound to adhere to the basic rights of the constitution. The *Supreme Court* explicitly confirmed this in 1833.[20] This situation meant that citizens were exclusively dependent on the constitutions of the various states for their protection against non-federal authorities. The American constitutional lawmaker reacted to this, after the civil war (1861-1865), by adding amendments 13, 14 and 15, and imposing several important human rights from the first ten amendments on the federal states.[21]

After an initial resistance to this extension of the federal constitution, the *Supreme Court* started to act on it and from then on accusations about alleged constitutional offences by the states were tested against the constitution. The Court, however, built-in an important reservation. By making use of the theory of selective incorporation only *certain* rights from the first ten amendments, the ones that were considered fundamental, were made compulsory for the American federal states.[22] Thanks to the *Warren Court* (1953-1969), we are generally led to believe that, in practice, an almost complete incorporation took place, but this is inaccurate.[23] Not all rights are fully incorporated. Substantive rights, like the freedom of speech, were quickly considered fundamental enough to be incorporated, but most procedural basic rights,[24] for example: from the fifth amendment , such as the right to silence,

just compensation'. For a definition of due process (a principle that ensures the right to an equal treatment before a court), of fair trial as a broader notion, and on the interrelation between 'fair trial' and 'due process', see Carsten Momsen & Marco Willumat, *Due Process and Fair Trial* in ELGAR ENCYCLOPEDIA OF CRIME AND CRIMINAL JUSTICE 102-113 (P Caeiro, S Gless, V Mitsilegas, Miguel J Costa, J de Snaijer, & G Theodorakakou, eds., Edward Elgar, 2024).

[20] Supreme Court, *Barron* v. *Baltimore* (Peters 1833, vol. 7) 24.

[21] For example, in the Fourteenth Amendment (1868), where the due process guarantee, as we know it, from the Fifth Amendment is made compulsory for state authorities. Cf Amendment 14, section 1: '*All persons born or naturalized in the United States, and subject to the jurisdiction thereof, are citizens of the United States and of the State wherein they reside. No State shall make or enforce any law which shall abridge the privileges or immunities of citizens of the United States; nor shall any State deprive any person of life, liberty, or property, without due process of law; nor to deny to any person within its jurisdiction the equal protection of the laws*'.

[22] See F Miatti, *La Due Process of Law américaine: quelle traduction française?* 74(2) REVUE DE DROIT INTERNATIONAL ET DE DROIT COMPARÉ 103 (1997); G Gunter, CONSTITUTIONAL LAW 405-585 (The Foundation Press 1985).

[23] *Id.* F Miatti, at 106.

[24] About the terms 'substantive' and 'procedural' rights: W Strasser, *The relationship between substantive rights and procedural rights guaranteed by the European Convention on Human Rights* in PROTECTING HUMAN RIGHTS: THE EUROPEAN DIMENSION. STUDIES IN HONOUR OF GÉRARD J. WIARDA 595-604 (F Matscher and H Petzold, eds., Karl Heymanns Verlag 1988).

were not incorporated or partially incorporated or only incorporated much later.[25] The most important items still not incorporated are the guarantee of a *grand jury* from the fifth amendment and the right to a jury in civil cases from the seventh amendment.[26]

The constitutional history of a legal system is made, in no small way, as a result of case law. Case law can lead to the creation of totally new constitutional rights or the extension of existing rights. It is common knowledge that the *Warren Court* has done revolutionary work in the field of constitutional renewal.[27] What comes to mind are its judgements on subjects, such as the freedom of speech, the right to privacy and the segregation of church and state.[28] There are also important judgements that can be pointed out in the field of procedural basic rights. In this way, the right to a lawyer in a defence case was broadly interpreted to make it also applicable to police interrogation.[29] The theory that illegally obtained evidence must be rejected from a case is also of judicial origin. The Court decided, in 1914, that evidence obtained in conflict with the fourth amendment- meaning illegally obtained – cannot be accepted as legitimate evidence.[30] Besides, in *Mapp* v. *Ohio* (1961),[31] the majority

[25] *Supra*, note 22 at 106. It is not our objective to specify precisely the size of the incorporation. Particularly the procedural rights from the *Bill of Rights* were omitted by the *Supreme Court* in their systematic incorporation and commitment, also after the easing or the incorporation criteria in the period after 1970.

[26] Only certain aspects of the jury guarantee for criminal cases were incorporated. The guarantee is only considered to be incorporated for serious criminal cases and consequently not for petty crimes. Neither is the requirement incorporated in the jury trial of criminal cases that the jury should consist of 12 members nor the requirement of a unanimous jury verdict. The jury administration of justice in civil cases, intended in the seventh amendment, is not considered to be incorporated. L Cavise, *Human Rights in the Trial Phase of the American System of Criminal Procedure* 8 NOUVELLES ETUDES PÉNALES *67* (1989) 82. For more details on civil rights that are not yet incorporated see: G. Bugh, INCORPORATION OF THE BILL OF RIGHTS: AN ACCOUNTING OF THE SUPREME COURT'S EXTENSION OF FEDERAL CIVIL LIBERTIES TO THE STATES 238 (Peter Lang Inc., 2022).

[27] About the 'Constitutional Revolution' in U.S. History due to the Warren Court: M Vitiello, I*ntroducing The Warren Court's Criminal Procedure Revolution: A 50-Year Retrospective*, 51 THE UNIVERSITY OF THE PACIFIC LAW REVIEW 621-632 (2020).

[28] T Koopmans, *The roots of judicial activism* in PROTECTING HUMAN RIGHTS: THE EUROPEAN DIMENSION. STUDIES IN HONOUR OF GÉRARD J. WIARDA 317-327 (F Matscher & H Petzold, eds., Karl Heymanns Verlag 1988).

[29] Supreme Court, *Escobedo* v. *Illinois*, United States Supreme Court Reports (US), Vol. 378, 1964, 478. Compare with judgement by the constitutional court in France: J Pradel, *Droit pénal comparé*, Dalloz, Paris 1995, 184.

[30] Supreme Court, *Weeks* v. *United States*, United States Supreme Court reports (US), Vol. 232, 1914, 383.

[31] '*We hold that all evidence obtained by searches and seizures in violation of the Constitution is, by that same authority, inadmissible in a state court. Since the Fourth Amendment's right of privacy*

Contd…

judged that this *exclusionary rule* should be considered as '*an essential part of the right to privacy*', whereupon it was decided to incorporate it.[32] In 1966, in *Miranda* v. *Arizona*, it was decided that, in addition to the already mentioned right to assistance of an advisor during police interrogation, it must be pointed out to suspects before they are interrogated by the police that they have the right to silence and the right to request legal assistance, otherwise statements received as evidence will be considered inadmissible.[33] The *Miranda* judgement illustrates the bandwidth of the exclusionary rule: from privacy to violations of the fifth and the sixth amendment.

Especially during the period of the *Burger Court* (1970-1986), a number of corrections were made to the acquis of the *Warren Court,* including the field of the '*Miranda-rules'* and the field of the *exclusionary rule,* that was connected with numerous exceptions.[34] The result of this case law is that some state courts in the US consider some 'new' human rights like those concerning the *exclusionary rule* not to be fundamental rights, but only procedural rules.[35]

has been declared enforceable against the States through the Due Process Clause of the Fourteenth, is enforceable against them by the same sanction of exclusion as is used against the Federal Government' (Supreme Court, *Mapp* v. *Ohio,* United States Supreme Court Reports (US), Vol. 367, 1961, 643; G Gunter, CONSTITUTIONAL LAW 437 (The Foundation Press 1985).

[32] This judgement followed twelve years after a judgement where the incorporation of the *exclusionary rule* was rejected on account of it not being fundamental enough. Cf Supreme Court, *Wolf* v. *Colorado*, United States Supreme Court Reports (US), Vol. 338, 1949, 25. The Court incorporated the fourth amendment in this case in the *due process clause* of the fourteenth amendment; but made an exception for the *exclusionary rule.* On the *exclusionary rule*: B George, *Due Process Rights of the Criminal Defendant in the Pre-Trial Phase* 8 NOUVELLES ETUDES PÉNALES 12 (1989) 26-28.

[33] Supreme Court, *Miranda* v. *Arizona,* United States Supreme Court Reports (US), Vol. 348, 1966, 436. Concerning the *Miranda rule: Supra*, note 32.

[34] C Blakesly, *The Role and Impact of Constitutionalism, Constitutional Courts & Supreme Courts on the Evolution & Development of Criminal Justice Systems: The Strange Trip in the U.S.* 17 NOUVELLES ETUDES PÉNALES 271 (1998) 279-300; C Whitebread, *The Burger Court's Counter-Revolution in Criminal Procedure* 24 WASHBURN LAW JOURNAL 470 (1985) 471-474. A summary of this last article is also included in Y Kamisar, W Lafave & J Israel, BASIC CRIMINAL PROCEDURE 116-118 (8th edn., West Publishing Co 1994). This step backwards leads to a plea to work less with the federal Constitution, and more with the fairly uncultivated ground of constitutional law in the federal states.

[35] B Latzer, STATE CONSTITUTIONS AND CRIMINAL JUSTICE 34-35 (Greenwood Press 1991).

III

Due Process: Transparency to Challenge AI Reliability and Trade Secrets

Protected under the fifth and the fourteenth amendments of the US constitution, the right to due process has been seen as an evolving concept granting judicial discretion in determining whether, given the particularities of a case, a law or practice is unfair.[36] Applying to the various stages of the criminal procedure (pre-trial, trial and post-trial), due process's general requirements include (according to the US Supreme Court): (a) fairness and reliability (*e.g.,* in investigations);[37] (b) respect for fundamental defence rights, such as the presumption of innocence,[38] the proof beyond reasonable doubt,[39] the right to cross-examination[40] or adversarial proceedings;[41] (c) the opportunity to be heard and present evidence;[42] (d) sufficient access to evidence;[43] (e) impartiality of judges.[44]

Aside from these general requirements, the right to due process can cover the protection of other rights.[45] Relevant, here, is the right to confrontation (sixth amendment), enabling the defence to effectively contest reliability of evidence.[46] To

[36] *See*, *Duncan* v. *Louisiana*, 391 US 145 (1968) para 168 ('*Duncan*').

[37] *See*, for instance, fairness and reliability in prosecution and identification by law enforcement: *Cone* v. *Bell*, 556 US 449 (2009) para 451; *Foster* v. *California*, 394 US 440 (1969) para 443.

[38] *Taylor* v. *Kentucky*, 436 US 478 (1978) para 490.

[39] *In re Winship*, 397 US 358 (1970) para 364; *Sandstrom* v. *Montana*, 442 US 510 (1979) para 520.

[40] *Chambers* v. *Mississippi*, 410 US 284 (1973) para 294.

[41] *Duncan*, para 187. Adversarial nature of proceedings is also promoted by the compulsory clause (under the Sixth Amendment of the US Constitution).

[42] *Mooney* v. *Holohan*, 294 US 103 (1935) paras 106-107.

[43] *Cone* v. *Bell*, 556 US 449 (2009) paras 451, 469-470.

[44] *Tumey* v. *Ohio*, 273 US 510 (1927) para 512; *Sheppard* v. *Maxwell*, 384 US 333 (1966) para 362.

[45] For instance, in the presentencing phase, due process can include protection of the right to be sentenced on the basis of accurate data and the right to individualised sentencing. See: *Loomis*, where the Supreme Court of Wisconsin favoured the consideration of a risk assessment tool for sentencing purposes and found no violation of these two fundamental rights, because, to that court, the defence could access and contest the tool's input and output and the risk assessment tool was considered for corroboration goals (it was not the sole/determinative element used to decide on the sentence).

[46] *Duncan*, paras 147-148; *Mattox* v. *United States*, 156 US 237 (1895) para 259; *Lee* v. *Illinois*, 476 US 530 (1986).

the US Supreme Court, this confrontation clause can hardly be reduced or watered down.[47]

Despite the acute need for reliability to be always checked, even where evidence appears obviously reliable,[48] the right to confrontation can in some cases be limited. One such case is the trade secret privilege, under which information that qualifies as a valid trade secret can be hidden from the defence.[49]

Misappropriation of trade secrets is, in principle, prohibited[50] when performed by improper means. The latter does not, however, include reverse engineering or lawful means of acquisition.[51] Lawful actions relating to trade secrets can also cover disclosures in judicial proceedings under certain circumstances (*e.g.,* confidentiality-safeguards).[52] Although the US Supreme Court has called for narrow interpretation

[47] *Coy* v. *Iowa*, 487 US 1012 (1988) paras 1020-1021 ('(…) The State suggests that the confrontation interest at stake here was outweighed by the necessity of protecting victims of sexual abuse. It is true that we have in the past indicated that rights conferred by the Confrontation Clause are not absolute and may give way to other important interests. The rights referred to in those cases, however, were not the right narrowly and explicitly set forth in the Clause, but rather rights that are, or were asserted to be, reasonably implicit -- namely, the right to cross-examine (…) the right to exclude out-of-court statements (…) and the asserted right to face-to-face confrontation at some point in the proceedings other than the trial itself (…) To hold that our determination of what implications are reasonable must take into account other important interests is not the same as holding that we can identify exceptions, in light of other important interests, to the irreducible literal meaning of the Clause: 'a right to meet *face to face* all those who appear and give evidence *at trial* (…) We leave for another day, however, the question whether any exceptions exist (…) Whatever they may be, they would surely be allowed only when necessary to further an important public policy (…)').

[48] *Melendez–Diaz* v. *Massachusetts*, 557 US 305 (2009) paras 317-318. *Crawford* v. *Washington*, 541 US 36 (2004) paras 61-62 ('(…) the Clause's goal is to ensure reliability of evidence, but it is a procedural rather than a substantive guarantee. It commands, not that evidence be reliable, but that reliability be assessed in a particular manner: by testing in the crucible of cross-examination (…) Dispensing with confrontation because testimony is obviously reliable is akin to dispensing with jury trial because a defendant is obviously guilty (…)').

[49] Under the 18 US Code (Chapter 90, §§ 1831-1839) on the protection of trade secrets, 'trade secret' is defined as information, whose holder has made reasonable efforts to keep it secret and whose economic value stems from its secrecy. See: 18 US Code § 1839, point (3).

[50] 18 US Code § 1839, point (5).

[51] 18 US Code § 1839, point (6).

[52] 18 US Code §§ 1833, 1835.

of privileges,[53] the trade secret privilege has been favoured in case law;[54] a favouring that, to some authors, can be particularly problematic in criminal cases.[55] Despite these critical developments in case law, it is noted that the US regime provides for procedural routes (*e.g.*, via protective orders)[56] that could permit conditional access to and disclosure of trade secrets (*e.g.*, under confidentiality-safeguards).[57]

In the artificial intelligence context, to meet the general requirements of due process, AI-made decisions and content would probably need to allow for enhanced accessibility and be subjected to rigorous scrutiny with a view to comprehending the way AI works and assessing reliability of a given output that would be considered by law enforcement or criminal justice authorities.[58]

This may not always be possible, especially considering some AI systems that are characterised by intrinsic opacity and complexity. Therefore, law enforcement and criminal justice authorities should avoid the use of AI, when making critical decisions that require adequate reason-giving and humanly comprehensible justifications that contemporary AI systems (and experts scrutinising them) may not be able to provide for. Such decisions could, for instance, include determinations on guilt or the severity of the sentence.[59] In other less critical implementations (*e.g.*, use of AI to merely enhance resolution of a low-quality CCTV-footage), AI could be

[53] *Pierce County* v. *Guillen*, 537 US 129 (2003) para 130 ('(…) Evidentiary privileges (…) must be construed narrowly because they impede the search for the truth (…)').

[54] A thorough analysis in: R Wexler, *Life, Liberty, and Trade Secrets: Intellectual Property in the Criminal Justice System* 70 STANFORD LAW REVIEW 1343 (2018).

[55] As Wexler points out, '[w]ithholding information from the accused because it is a trade secret mischaracterizes defence advocacy as a business competition' (R Wexler (n 54) 1429). See also 1401-1402 ('(…) Overvaluing trade secret claims in criminal cases is inconsistent with principles of procedural justice (…) the trade secret privilege's balancing test is suspect because it appears to place pure financial interests on par with life and liberty (…)').

[56] Federal Rule of Criminal Procedure 16 on discovery and inspection.

[57] Concerning protective orders and other legal instruments that could allow for access without violating trade secrecy, see: *Supra,* note 54 at 1409ff; J Villasenor & V Foggo, *Artificial Intelligence, Due Process, and Criminal Sentencing* 2020 MICHIGAN STATE LAW REVIEW 295 (2020) 343ff.

[58] A critical analysis of black-box technologies and the need for recognising the right to interpretable ('glass-box') AI in: Br Garrett & C Rudin, *The Right to a Glass Box: Rethinking the Use of Artificial Intelligence in Criminal Justice* 109(3) CORNELL LAW REVIEW 561 (2023).

[59] See, for instance, rules on whether and how to consider risk assessment technologies for sentencing purposes: 2024 California Rules of Court, 'Standard 4.35. Court use of risk/needs assessments at sentencing', point (e) ('(…) (1) The results of a risk/needs assessment should not be used to determine: (A) Whether to incarcerate a defendant; or (B) The severity of the sentence (…)') *available at*: https://www.courts.ca.gov/cms/rules/index.cfm?title=standards&linkid=standard4_35 (last visited Sep. 08, 2024).

considered, yet, in our view, solely for corroboration purposes and provided that, first, materials, other than AI, can independently support a judicial or other (*e.g.,* law enforcement) decision and, second, transparency and other essential safeguards are in place[60] (*e.g.,* to ensure high-levelled accessibility and reviewability of AI-tools). In any event, if regulators opted for the use of IP-protected AI, legal instruments (like the above-mentioned protective orders) could allow for conditional access (*e.g.,* under confidentiality-safeguards);[61] and, alternatively, if such conditional access were (perhaps unreasonably) not granted, some optimal levels of transparency and accessibility could to a certain degree be met by disclosing relevant materials that are not protected by IP (such as ancillary information on methodologies, statistics or training data).[62]

Coming back to the previous paragraph (see right above), where we argued for a prohibition of black box AI 'when making critical decisions' in criminal prosecution. Is there a reason for the justice apparatus to work with such black box machines? Duke Professor Brandon L Garrett, in his interventions and writings, famously goes against the current and holds that the need for black box AI systems is overstated, many interpretable AI systems perform as well and have clear advantages in terms of due process rights.[63] To the author, the burden rests on the government to justify any departure by a law enforcement authority (or the judge) 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. Along with his co-author, he calls for national and local case law and regulation to protect the right to glass-box AI in all criminal cases.

[60] See, by analogy, *Loomis*, where the Supreme Court of Wisconsin permitted the consideration of a risk assessment technology for corroboration purposes and conditioned it upon disclaimers (including the provision of information on opacity, bias or the need for review and validation).

[61] See, in this regard: *Supra* note 54, R Wexler 1409ff; *Supra* note 57, J Villasenor & V Foggo 343ff.

[62] A discussion in: L Chan, *The Weaponization of Trade Secret Law* 124 COLUMBIA LAW REVIEW 703 (2024), arguing that algorithms may be protectable as trade secrets; albeit ancillary information can be made visible with a view to promoting fairness without unduly limiting trade secrecy protection.

[63] Br Garrett, *Rule 702 Amendments and their Impact on Admissibility of AI 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 Nov. 12, 2024); Brandon Garrett and C Rudin, *Interpretable algorithmic forensics* 120(41) PROC NATL ACAD SCI US (2023); *Supra* note 58, Br Garrett & C Rudin, 561. By 'interpretable' AI, the authors refer to predictive models whose calculations are inherently understandable. By 'explainable', they refer to a system that provides a post hoc explanation for its model, which could be a black box model (page 572).

IV

Requirements of Evidence Law (the US Federal Rules of Evidence)

It can be generally stated that every type of criminal procedure in Western legal systems is based on either an accusatory (common law) or an inquisitorial (civil law) procedural model. The US is an example of the former model; 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 law 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 any 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:[64]

a) within the accusatory tradition, the belief does not exist that the objective truth lies in evidential materials. Both parties have to 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;

b) 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

[64] M Caianiello, *Adversarial and Inquisitorial Criminal Procedure* in ELGAR ENCYCLOPEDIA OF CRIME AND CRIMINAL JUSTICE 46-62 (P Caeiro, *et al.*, eds., Edward Elgar, 2024); T Decaigny & P De Hert, *You can change the color of your hair, not your hair. Converging is what inquisitorial and adversarial systems rarely do* in VEELZIJDIGE GEDACHTEN. LIBER AMICORUM CHRISJE BRANTS 235-244 (C Kelk, *et al.*, eds., Den Haag, Boom Lemma Uitgevers, 2013); T Decaigny, *Inquisitorial and Adversarial Expert Examinations in the Case Law of the European Court of Human Rights* 5(2*)* NEW JOURNAL OF EUROPEAN CRIMINAL LAW (2014) 149-166; Fr Tulkens, *Criminal Procedure: Main Comparable Features of the National Systems* in THE CRIMINAL PROCESS AND HUMAN RIGHTS 8-9 (M Delmas-Marty, ed., Martinus Nijhoff Publishers, 1995). Th Thaman, *Trial by Jury and the Constitutional Rights of the Accused in Russia* 4(1) EAST EUROPEAN CONSTITUTIONAL REVIEW 77 (1995) 77-80. For a more legal theoretical account of the difference in legal culture, see: B Edelman, *Universality and Human Rights* in THE CRIMINAL PROCESS AND HUMAN RIGHTS 97-107 (M Delmas-Marty, ed., Martinus Nijhoff Publishers 1995); F Macchiarola, *Finding the Truth in an American Criminal Trial: Some Observations* in PROCEEDINGS OF THE FIRST WORLD CONFERENCE ON NEW TRENDS IN CRIMINAL INVESTIGATION AND EVIDENCE 85-87 (J Nijboer & J Reijntjes, eds., Koninklijke Vermande 1997).

themselves, to abandon the law suit and to agree upon a certain result through negotiations;[65]

c) 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);
d) the collection of evidential materials is the complete responsibility of the parties themselves, and the accused, who is considered a full party; and she can decide him/herself whether she wants to actively contribute to 'her' truth (process autonomy);
e) coercive measures can only be employed with the authority of the judge.

Different to this accusatory procedure, the inquisitory procedure draws from the idea of an objective, material truth that 'must' be found through a process where a professional judge forms the pivotal figure. 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 investigations where the exercise of collecting evidence is done by the public prosecution's department or by an independent magistrate (the examining magistrate). The department of public prosecution is involved in both the investigation and the prosecution. 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 the examining magistrate on her own initiative, without a formal request from the prosecuting party.

It is useful to remind readers of these aspects of comparative criminal procedure. As Professor Gless has suggested, this background has an important impact on AI in court. Evidence that is considered unreliable or illegal will not make it into court but excluded in the American systems based on the jury system. In the inquisitorial systems that work with bench judges, *all* the evidence will be seen by them and then will have to justify afterwards their use of it (if they use it).[66]

[65] 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.

[66] S Gless, *Rules on Expert Testimony based on a Comparative Perspective on Device 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 Nov. 12, 2024).

The US has a system of state criminal courts and agencies and a system of federal criminal courts and agencies. Criminal procedural law is partly governed by state law and state constitutions and by federal law and the federal Constitution (see *above*). The Federal Rules of Evidence govern or influence the admission or exclusion of evidence in most proceedings in the US courts.[67] These Federal Rules of Evidence became federal law on January 2, 1975, when President Ford signed the *Act to Establish Rules of Evidence for Certain Courts and Proceedings*.[68] As enacted, the Evidence Rules included amendments by Congress to the rules originally proposed by the Supreme Court. The Evidence Rules were last amended in 2023.[69]

According to the US Federal Rules of Evidence, information is admissible, if it is 'relevant'[70] *i.e.,* having a tendency to prove a fact.[71] Relevant materials can be excluded, if their probative value can be substantially outweighed by, among others, unfair prejudice or misleading.[72] In general, judges are given discretion, acting as gatekeepers, for adequately reliable evidence; this is not the case in continental systems.[73]

The previous paragraphs on evidence can help law enforcement and criminal justice authorities to approach AI-made decisions and content *via* a safe route. AI's output can be deemed 'relevant' as required by the Federal Rules, insofar as it can prove something, *e.g.,* that something happened. For instance, an AI tool that enhances or augments a video by adding content to a low-quality footage or significantly modifying it to increase quality, details or resolution, may not be relevant, if it creates new content and fails to show what really happened.[74] Furthermore, AI-

[67] The Supreme Court submitted proposed Federal Rules of Evidence to Congress on February 5, 1973, but Congress exercised its power under the Rules Enabling Act to suspend their implementation.

[68] Act to Establish Rules of Evidence for Certain Courts and Proceedings, *Pub L* No 93-595.

[69] See full text *available at*: https://www.uscourts.gov/rules-policies/current-rules-practice-procedure/federal-rules-evidence (last visited Nov. 12, 2024).

[70] Federal Rule of Evidence 402. See also Federal Rule of Evidence 401.

[71] Notes of Advisory Committee on Federal Rule of Evidence 401, citing California Evidence Code §210 that in turn defines 'relevant evidence' in terms of its tendency to prove a fact.

[72] Federal Rule of Evidence 403.

[73] S Gless, Rules on Expert Testimony based on a Comparative Perspective on Device 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 Nov. 12, 2024), referring to the lack of a Rule 403-like provision in the Swiss, German and Dutch Criminal Procedure ('(…) as the bench is the trier of facts and judges are expected to be 'professionals', there is an 'open gate' policy in inquisitorial tradition (…)').

[74] This was the case in *Puloka*, excluding an AI-augmented video that had used unintelligible methods and failed to reveal what really happened. See: *Puloka*, para 13

Contd…

related evidence should be excluded, if its probative value is substantially outweighed by, among others, potential unfair prejudice, stemming from possible error-rates, bias or lack of adequate testing. Such an exclusion could be justified, where, for example, law enforcement authorities use opaque, possibly untested and biased AI-instruments for identification or other similar goals.[75]

In her recent speech, the Director of the Fourth Amendment Centre at the National Association of Criminal Defence Lawyers, Jumana Musa, focused on the due process requirement that the State must provide defendants with all evidence in its possession, that is, material to 'either guilt or punishment, irrespective of the good faith or bad faith of the prosecution'.[76] In the AI context, this could raise critical questions: what is the technology? how is it used by an entity (*e.g.,* law enforcement)? how was it used in a particular case? has it been validated (*e.g.,* for a state's own population)?[77] Things may become more troublesome in light of the need that the defence justifies its discovery request: what is a given defendant endeavouring to demonstrate (*e.g.,* unreliability of AI? misuse or improper application of AI? erroneous interpretation of AI's output? lack of expertise on the part of the AI-analyst?)?[78]

Of relevance here is the recent *Arteaga*-case, where the appellate court favored the defense concerning discovery for an opaque, possibly untested and biased, face recognition technology.[79] Although *Arteaga* can be seen as a milestone on AI-related

('(…) The video evidence produced by the Topaz Video AI enhancement model does not satisfy ER 401, as the resulting video does not show with integrity what happened but uses opaque methods to represent what the AI model 'thinks' should be shown (…)').

[75] In this regard, see: *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. See also: ACLU New Jersey, *New Jersey Appellate Division one of first courts in country to rule on constitutional rights related to facial recognition technologies* (Jun. 8, 2023) *available at*: https://www.aclu-nj.org/en/press-releases/new-jersey-appellate-division-one-first-courts-country-rule-constitutional-rights (last visited Sep. 08, 2024). On the risk of unfair prejudice outweighing relevance, see *Puloka*, para 14 ('(…) admission of this AI-enhanced evidence would lead to a confusion (...) and could lead to a time-consuming trial within a trial about the non-peer-reviewable-process used by the AI model, such that any relevance is substantially outweighed by the danger of unfair prejudice under ER 403 (…)').

[76] 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 Nov. 12, 2024) referring to *Brady* v. *Maryland*, 373 US 83 (1963) 87.

[77] J Musa (n 76).

[78] J Musa (n 76).

[79] *State* v. *Arteaga* (Superior Court of New Jersey, Appellate Division; Docket No A-3078-21;

Contd…

discovery requests, fast tech-developments may render this case already passé. What seems clear is that the defence could rely upon and follow other routes to scrutinise AI. For instance, as Musa recommends, public records requests, based on the Freedom of Information Act or state laws with a similar purpose, could be used by the defence to find more about the logic of an AI system.[80] Still, even such requests may offer no useful insights into AI, given that state authorities using complex technologies (like automated licence plate readers, ALPR) can enjoy unfettered discretion in gathering huge amounts of data or analysing patterns in an unintelligible manner (*e.g.*, with a view to establishing suspicion or making hot lists).[81]

V

Requirements of Expert Evidence (Under the US Federal Rules of Evidence)

Explaining the exact logic behind each individual decision might not always be feasible in the case of AI. Moreover, courts themselves might be hesitant to provide detailed insights into how these systems function.[82] For this reason, as Gless, Lederer and Weigend note, the way AI-driven evidence is presented in courts is critical for ensuring its comprehensibility and trustworthiness.[83] Even experts may

Decided on 7 June 2023). In this case, the key question was whether a defendant who was identified using a facial recognition system is entitled to a detailed discovery on the system and the specifics of how he was identified).

[80] *Supra* note 76.

[81] *Supra* note 76.

[82] For example, in case C/09/662309 / HA RK 24-104 before the District Court of The Hague, the court voiced concerns that disclosing how an AI-based transaction-monitoring system operated might reveal vulnerabilities that could be exploited by malicious actors. The Court's ruling emphasised that revealing too much could compromise the system's ability to prevent crime. However, this approach seems overly cautious, prioritising system security over the transparency needed to uphold defence rights.

[83] S Gless, Fr Lederer & Th Weigend, *AI-Based Evidence in Criminal Trials?* 59(1) TULSA LAW REVIEW 1 (2024) 34-35, through a comparative analysis of the approach to device evidence between the US and Germany, the authors conclude on a procedural solution based on expert testimony in German criminal procedure, which allows for neutral expert involvement to ensure the comprehensibility and trustworthiness of AI-driven evidence in court. The authors also explore a technological solution (pages 31 to 34) involving standards for reliability, validation and certification, suggesting the potential for AI-driven verification tools, such as Artificial Counter-Intelligence (ACI), to assess accuracy and integrity of AI-driven evidence. While promising, one must wait for AI's further

Contd…

struggle to clarify how an AI-driven device assesses human behaviour or to establish a clear causal link between input data and resulting conclusions.

A considerable part of the US Federal Rules of Evidence is devoted to the question of what is 'scientific knowledge' and who can qualify as an expert in court. On expert evidence, a witness can testify, if: she is qualified (*e.g.*, in terms of experience);[84] her expertise can assist the court in comprehending evidence or deciding on a matter; and her testimony has been based on 'sufficient facts or data'[85] and resulted from reliable principles and methods that have been reliably applied to the factual particularities of a concrete case.[86] Settled case law, following *Frye*,[87] *Kelly*,[88] and *Daubert*,[89] has set out some minimum elements to be taken into account,

incorporation into criminal proceedings to evaluate ACI's potential in vetting such evidence.

[84] Experience as such can suffice for the expert to be qualified under Federal Rule of Evidence 702. Where experts use as their basis only or mainly their experience, they need show: the way in which experience results in their conclusions; the reasons why their experience constitutes an adequate ground for the testimony; and the way in which their experience is reliably applied to the case at hand. See: Federal Rule of Evidence 702, Committee Notes on Rules – 2000 Amendment, with further references to case law.

[85] The requirement that the testimony rely upon sufficient data can refer to reliable opinions of other experts. See: Federal Rule of Evidence 702, Committee Notes on Rules – 2000 Amendment, with further references to case law.

[86] Federal Rule of Evidence 702.

[87] Before the adoption of the Federal Rules of Evidence, courts applied the *Frye*-test. Coming from a 1923 judgement of the Supreme Court of the District of Columbia, this test demands that evidence be admitted, if it is generally accepted as reliable by the relevant scientific community. *Frye* v. *United States*, 293 F 1013 (DC Cir 1923).

[88] *Kelly* comes from a 1976 decision of the Supreme Court of California on proof of reliability, where new techniques were involved. *Kelly* referred to three requirements: general acceptance; adequate qualification of the expert; and application of proper procedures to the case at hand (*People* v. *Kelly*, 17 Cal 3d 24 (Cal 1976) para 30). That court concluded that *Frye* was the test to be relied upon for evaluating expert evidence, which, by that time, had not gained general acceptance (*Kelly*, paras 32, 40-41). Case law on *Kelly* and *Frye* has stressed that these tests are primarily focused on the requirement of 'general acceptance' and mainly aimed at not having the factfinder misled by 'unproven and ultimately unsound' techniques. See, among others: *People* v. *Therrian*, 113 Cal App 4th 609 (Cal Ct App 2003) paras 614-616.

[89] *Daubert* is a 1993 judgement of the US Supreme Court, holding that the *Frye*-test was replaced by the implementation of the Federal Rules of Evidence (*Daubert* v. *Merrell Dow Pharmaceuticals Inc*, 509 US 579 (1993) para 587; in this case, the plaintiff party, the parents of two minors, argued that the mother's ingestion of a drug (Benedictin) caused the children to have birth defects). As courts have clarified, *Daubert* made the *Frye*-test compliant with the Federal Rules of Evidence, especially Rule 702 on expert evidence, by rendering its 'general acceptance' requirement one among the factors to be considered for

Contd…

when determining reliability and relevance of expert evidence: *testability* (whether evidential materials can be or have been sufficiently tested); *peer-review and publications* (whether evidential materials have been subjected to sufficient peer-reviewing and publications); *possible error-rates* (the known or potential rate of error); *standards controlling a methodology's operation* (whether standards on a methodology's operation are in place and maintained); and *acceptability* by the relevant scientific community.[90]

The Federal Rules on expert-evidence (discussed *above*) require at least three things when it comes to experts confronted with AI. *First*, human experts, scrutinising the AI's input, processing and output, should be adequately qualified, for instance, in terms of technical skills or experience and education on AI's complex analysis. *Second*, these experts must be capable of assisting in the fact-finding process by enabling the court to comprehend the AI-related evidence or decide on a matter (*e.g.*, by explaining that an AI-enhanced video only improves resolution of a low-quality footage and in no way adds content to it).[91] *Third*, they should base their testimony on adequate data, as well as follow reliable principles and methods that are, in addition, reliably applied to the particularities of the case (*e.g.*, adequate data could require big training datasets that would, moreover, be accurate and highly representative).

In addition to these three requirements, *Daubert* (*above*) demands sufficient testing, peer-review and publications, low/no error-rates, as well as general acceptance by the AI-community;[92] a community that, given the ever-growing sophistication of AI, may not yet have fully comprehended and reached consensus on numerous complex processing operations.[93]

admissibility and by enlisting some minimum key factors that need be addressed (*Daubert*, paras 588, 593-594).

[90] An analysis of these minimum elements in: Committee Notes on Rule 702 (2000, 2011 and 2023 Amendments).

[91] See *Puloka*, where the AI-enhanced video, adding its own content via opaque methods, was excluded.

[92] For a discussion on AI-related evidence and the *Daubert*-standard, see: S Gless, *AI in the Courtroom: A Comparative Analysis of Machine Evidence in Criminal Trials* 51 GEORGETOWN JOURNAL OF INTERNATIONAL LAW (2020) 195, 244ff; Patrick W Nutter, *Machine Learning Evidence: Admissibility and Weight* 21 JOURNAL OF CONSTITUTIONAL LAW 919 (2019) 931ff.

[93] See *Puloka*, where the opaque methods, used to enhance the video, had been neither sufficiently tested/peer-reviewed nor generally accepted by the scientific community (para 10: '(…) The Topaz Video AI enhancement tool(s), which utilize 'machine learning' algorithms, have not been peer-reviewed by the forensic video analysis community, are – at the present time – not reproducible by that community, and are not accepted generally in that community (…)').

Both Sabine Gless[94] and Jumana Musa[95] are critical toward testing expert evidence in court in the context of AI. We are admittedly not there yet, given the absence of forensic standards, external validation, industry standards or testing in real world conditions.[96]

VI

The 2023 Amendment to the Federal Rule of Evidence 702

Federal Rule of Evidence 702 is the main rule on expert evidence that federal (and most state) courts rely upon when determining whether someone can qualify as an expert. This rule was amended in 2023. The amended text stipulates that:

> '(…) A witness who is qualified as an expert by knowledge, skill, experience, training, or education may testify in the form of an opinion or otherwise *if the proponent demonstrates to the court that it is more likely than not that*:
> (a) the expert's scientific, technical, or other specialized knowledge will help the trier of fact to understand the evidence or to determine a fact in issue;
> (b) the testimony is based on sufficient facts or data;
> (c) the testimony is the product of reliable principles and methods; and
> (d) *the expert's opinion reflects a reliable application of the principles and methods to the facts of the case* (…)' (emphasis added).

The above phrasing '*if the proponent demonstrates to the court that it is more likely than not that*' was added; and point (d), ('*the expert's opinion reflects a reliable application of the principles and methods to the facts of the case*') replaced the old text ('*the expert has reliably applied the principles and methods to the facts of the case*').[97] On the one hand, in the past, many judges presumed that the prosecutor's expert is reliable; hence, they imposed no burden on the prosecutor to show that '*it is more likely than not that*', among others, the expert can help the court in comprehending/deciding on a matter or the testimony relies upon adequate data and has resulted from reliable principles/methods. This will now change. On the other hand, the second amendment suggests that it is not enough that a method is reliable; the conclusion of the expert must also be reliable, reflecting '*a reliable application of the principles and methods to the facts of the case*'. It has traditionally been assumed that the expert's

[94] *Supra* note 73.

[95] *Supra* note 76.

[96] *Supra* note 76.

[97] See Communication from the Chief Justice, the Supreme Court of the United States transmitting Amendments to the Federal Rules of Evidence that has been adopted by the Supreme Court, pursuant to 28 USC 2072 (118th Congress, 1st Session, House Document 118–33) *available at*: https://www.govinfo.gov/content/pkg/CDOC-118hdoc33/pdf/CDOC-118hdoc33.pdf (last visited Nov. 12, 2024).

conclusion is not checked, if the method is reliable, even where the expert might 'exaggerate'. Under the amended provision, an expert's opinion must now be scrutinised by someone other than the expert (benefiting the defence who no longer carries the burden).[98]

VII

Demands of Equal Protection: Strict-Scrutiny of AI-Usage Interfering with Fundamental Rights

Protected under the Fourteenth Amendment of the US Constitution, the right to equal protection[99] prohibits discrimination that is unfair (lacking reasonable grounds and resulting in arbitrariness).[100] There are various tests and levels of scrutiny that courts can apply to ascertain unfair discrimination.[101] Three basic types of judicial scrutiny are strict scrutiny, intermediate scrutiny and rational-basis scrutiny.

More precisely, strict scrutiny is often applied to cases, involving possible violation of fundamental rights or suspect classification;[102] that is, discrimination on the basis

[98] A discussion on the two amendments in: Brandon Garret, *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 Nov. 12, 2024).

[99] For a discussion on the principles of equality and non-discrimination, as well as a thorough analysis of the right to equal protection and relevant case law, see: European Parliamentary Research Service, *The principles of equality and non-discrimination - A comparative law perspective - United States of America* (EPRS, March 2021) 50ff *available at*: https://www.europarl.europa.eu/RegData/etudes/STUD/2021/689375/EPRS_STU(2021)689375_EN.pdf (last visited Sep. 08, 2024).

[100] See, among others: *Lindsley* v. *Natural Carbonic Gas Co*, 220 US 61 (1911) paras 81-82.

[101] For the various levels of judicial scrutiny and relevant case law, see: R R Kelso, Standards of Review under the Equal Protection Clause and Related Constitutional Doctrines Protecting Individual Rights: The 'Base Plus Six' Model and Modern Supreme Court Practice 4(2) JOURNAL OF CONSTITUTIONAL LAW 225 (2002). See also: G B Daniels & R Pereira, *Equal Protection as a Vehicle for Equal Access and Opportunity: Constance Baker Motley and the Fourteenth Amendment in Education Cases* 117(7) COLUMBIA LAW REVIEW 1779 (2017).

[102] On strict scrutiny, see, among others: A K Blair, *Constitutional Equal Protection, Strict Scrutiny, and the Politics of Marriage Law* 47(4) CATHOLIC UNIVERSITY LAW REVIEW 1231 (1998); R G Spece & D Yokum, *Scrutinizing Strict Scrutiny* 40 VERMONT LAW REVIEW 285 (2015); Ev Gerstmann & Chr Shortell, *The Many Faces of Strict Scrutiny: How the Supreme Court Changes the Rules in Race Cases* 72(1) UNIVERSITY OF PITTSBURGH LAW REVIEW 1 (2010).

of race, national origin, religion or alienage, as well as discrimination against 'discrete and insular' minorities.[103] For instance, on racial discrimination, the US Supreme Court, in principle, applies a stringent test,[104] requiring a heavy burden of justification and necessity of the measure or law to achieve the goal pursued (the compelling state interest).[105] Under the intermediate scrutiny-test, often conducted in gender-discrimination cases,[106] the discriminating law or practice must, among others, serve an important goal (a government interest) by means closely linked (proportionate) to that goal.[107] Lastly, the rational basis-scrutiny, applied to cases involving no suspect classification,[108] can require rational linkages between the discriminatory measure or law and the goal pursued.[109]

It is noted that, to establish undue discrimination, the US Supreme Court can demand that the defence demonstrate not only discriminatory impact (effect), but also discriminative intent[110] or purpose.[111]

In the AI context, the *above* could mean that, depending on the particularities of a given case, AI-made decisions and content would be subjected either to rigorous

[103] On suspect classification and the concept of 'discrete and insular minorities', see: M Strauss, *Reevaluating Suspect Classifications* 35 SEATTLE UNIVERSITY LAW REVIEW 135 (2011); J Harras, *Suspicious Suspect Classes - Are Non-Immigrants Entitled to Strict Scrutiny Review under the Equal Protection Clause?: An Analysis of Dandamudi and LeClerc* 88(3) ST JOHN'S LAW REVIEW 849 (2014).

[104] *Korematsu* v. *United States*, 323 US 214 (1944) para 216; *Yick Wo* v. *Hopkins*, 118 US 356 (1886) para 374.

[105] *McLaughlin* v. *Florida*, 379 US 184 (1964) paras 194, 196; *Loving* v. *Virginia*, 388 US 1 (1967) para 11.

[106] On intermediate scrutiny, see: R Holoszyc-Pimentel, *Reconciling Rational-Basis Review: When Does Rational Basis Bite?* 90 NEW YORK UNIVERSITY LAW REVIEW 2070 (2015); A Bhagwat, *The Test That Ate Everything: Intermediate Scrutiny in First Amendment Jurisprudence* 2007(3) UNIVERSITY OF ILLINOIS LAW REVIEW 783 (2007); *Let the End Be Legitimate: Questioning the Value of Heightened Scrutiny's Compelling- and Important-Interest Inquiries* 129(5) HARVARD LAW REVIEW 1406 (2016).

[107] *Craig* v. *Boren*, 429 US 190 (1976) para 197.

[108] On the rational basis scrutiny, see: K R Eyer, *The Canon of Rational Basis Review* 93(3) NOTRE DAME LAW REVIEW 1317 (2018); *Supra* note 106.

[109] For discrimination based on wealth/poverty, see: *McDonald* v. *Board of Election*, 394 US 802 (1969) para 807; *Bullock* v. *Carter*, 405 US 134 (1972) para 144; *San Antonio School District* v. *Rodriguez*, 411 US 1 (1973) (*Rodriguez*) para 44; *Maher* v. *Roe*, 432 US 464 (1977). On age discrimination, see: *Massachusetts Bd of Retirement* v. *Murgia*, 427 US 307 (1976); *Vance* v. *Bradley*, 440 US 93 (1979) para 109; *Gregory* v. *Ashcroft*, 501 US 452 (1991) para 453.

[110] This has been the case with race and gender discrimination. See: *Washington* v. *Davis*, 426 US 229 (1976) para 241; *Mobile* v. *Bolden*, 446 US 55 (1980) paras 66, 70; *Personnel Administrator of Mass* v. *Feeney*, 442 US 256 (1979) para 274.

[111] A discussion on discriminatory intent/purpose and relevant case law, in: R W Galloway, *Basic Equal Protection Analysis* 29 *SANTA CLARA LAW REVIEW* 121 (1989) 131ff.

tests, *e.g.*, requiring a heavy burden of justification and necessity of the discriminating practice, or to more relaxed checks, focusing on the importance of the objective pursued or the rational linkages between the discriminating practice and the goal served. This approach could result in applying stringent scrutiny to critical practices, such as AI-uses by law enforcement leading to unfair discrimination against vulnerable minorities.[112] However, there might be a chance that judges apply weaker tests to cases, involving the processing of seemingly less/noncritical data (like age) as inputs of AI-made decisions and contents, whose implementation could nevertheless interfere with fundamental rights.[113]

The choice of judges for a specific scrutiny test has not crystalised yet in the AI context. But, given that strict scrutiny is the test applicable to cases involving possible violation of fundamental rights, such strictness should, in our view, be used to assess (un)fairness of all AI-made decisions and contents, whose implementations can interfere with human rights and freedoms.

More problematic, in the future, might be the requirement to demonstrate discriminatory intent/purpose. This could impose a particularly heavy and perhaps unreasonable burden on a given defendant to show that, for instance, a concrete factor, considered by an AI-tool, was included for the goal of discriminating against her and not to promote statistical precision or accuracy.[114] In the AI context, where a tool's unintelligible processing may, regardless of its designer's intent, result in

[112] On bias in predictive policing, see, among others: R Richardson, J M Schultz & K Crawford, *Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice* 94 NEW YORK UNIVERSITY LAW REVIEW 192 (2019). See also: Will Douglas, *Predictive policing algorithms are racist. They need to be dismantled*, TECHNOLOGY REVIEW (Jul. 17, 2020) *available at*: https://www.technologyreview.com/2020/07/17/1005396/predictive-policing-algorithms-racist-dismantled-machine-learning-bias-criminal-justice/ (last visited Sep. 08, 2024).

[113] In this regard, see: EDPB, *Report of the work undertaken by the ChatGPT Taskforce* (European Data Protection Board, 23 May 2024) *available at*: https://www.edpb.europa.eu/system/files/2024-05/edpb_20240523_report_chatgpt_taskforce_en.pdf (last visited Sep. 08, 2024). See also: Anna Bacciarelli and Paul Aufiero, *Pandora's Box: Generative AI Companies, ChatGPT, and Human Rights: What's at Stake in Tech's Newest Race?*, HUMAN RIGHTS WATCH (May 3, 2023), referring to the use of ChatGPT and similar technologies: '(…) Even when we enter seemingly mundane information into generative AI search or chatbots, this could be used to build a picture of who we are (…)' *available at*: https://www.hrw.org/news/2023/05/03/pandoras-box-generative-ai-companies-chatgpt-and-human-rights (last visited Sep. 08, 2024).

[114] See, by analogy, *Loomis*, where 'gender' was seen as an accuracy-enhancing factor, whose inclusion in the risk assessment had not (to the court) been aimed at unfairly discriminating against defendants. See also: *People* v. *Osman*, H037818 (Cal Ct App, 8 April 2013), adopting a similar (accuracy-enhancing) approach to the inclusion (in the risk assessment) of the factor of 'co-living with a partner' prior to marriage.

unfairly discriminating against certain groups,[115] proof of discriminative impact could suffice to establish undue discrimination.

VIII

Demands of Legality and Privacy: *Ex-ante* and Brightline Rulemaking on Data Processing

The principle of legality in criminal law is firmly established in American law: crimes cannot be introduced without proper laws. The same principle together with the principle of certainty and the principles of data privacy law explain why the idea of a legal basis, that is detailed enough, also govern the framing of regulations for police and law enforcement authorities in general. Law enforcement and criminal justice authorities must admittedly have some minimum access to personal information to better investigate or adjudicate on a case. But this access must, however, be adequately regulated and, under the circumstances, limited with a view to protecting the right to privacy and the protection of personal data, as well as guaranteeing legal certainty, foreseeability, integrity, security, reliability or accuracy of such data, whose processing may have an impact on ascertaining the truth during the fact-finding or other processes.

The extent to which privacy is protected under the US Constitution remains unclear.[116] What appears to be clear is that the US have long followed a piece-meal approach, with specific legal instruments targeted at concrete technological implementations.[117] As we have argued elsewhere, this approach has enabled the US to effectively protect the right to privacy and the protection of personal data

[115] See, for example: Allen Smith, *AI: Discriminatory Data In, Discrimination Out*, SHRM (Dec. 11, 2019) '(…) It is unlikely that an AI-enabled software would be intentionally developed to discriminate against minorities or women (…) the larger risk is that these tools may unintentionally discriminate against a protected group (…)' *available at*: https://www.shrm.org/topics-tools/employment-law-compliance/ai-discriminatory-data-discrimination (last visited Sep. 08, 2024).

[116] Some dimensions of the right to privacy may be protected under the First Amendment (on privacy of beliefs) or the Fourth Amendment (on searches and seizures).

[117] See, among others: Fair Credit Reporting Act, 15 USC § 1681; Electronic Communications Privacy Act, 18 USC § 2510; Cable Communications Policy Act, 47 USC § 551; Video Privacy Act, 18 USC § 2710. More examples in: G Zanfir-Fortuna, *America's 'Privacy Renaissance': What to Expect under a New Presidency and Congress: A Deep Dive into US Privacy Legislation and Implications for US-EU-UK Relations*, ADA LOVELACE INSTITUTE (Dec. 17, 2020) *available at*: https://www.adalovelaceinstitute.org/blog/americas-privacy-renaissance/ (last visited Sep. 08, 2024).

against intrusive practices *via* brightline laws from various areas (including criminal law), as well as through various modes of regulation, such as the market.[118]

This also seems to be the case with AI/personal data concerns in the criminal justice system. By way of example, we refer below to two recent legal instruments (at federal and state level) that, though not expressly aimed at protecting privacy, seem to impose rigorous obligations and set out concrete safeguards, ensuring high-levelled transparency, reliability or accuracy of data processing operations.

At federal level, the 2024 AI Policy, under the title 'Memorandum for the Heads of Executive Departments and Agencies' is aimed at advancing governance, innovation and risk management in the use of AI by federal entities.[119] Setting out a rigorous baseline for responsible AI-implementations, it imposes concrete obligations on various authorities, including law enforcement agencies. Its requirements range from transparency obligations (for instance, on risk assessments)[120] to independent scrutiny (such as impact assessments concerning

[118] P De Hert & G Bouchagiar, *Facial Recognition, Visual and Biometric Data in the US: Recent, Promising Developments to Regulate Intrusive Technologies* 7(29) BRUSSELS PRIVACY HUB WORKING PAPER (Oct. 2021) *available at*: https://brusselsprivacyhub.eu/publications/BPH-Working-Paper-VOL7-N29.pdf. See also: P De Hert & G Bouchagiar, *European Biometric Surveillance, Concrete Rules and Uniform Enforcement. Beyond Regulatory Abstraction and Local Enforcement* in CAMBRIDGE HANDBOOK ON FACIAL RECOGNITION IN THE MODERN STATE 139 (R Matulionyte & M Zalnieriute, eds., Cambridge University Press 2024).

[119] This policy was published by the Office of Management and Budget (OMB) of the Executive Office of the President on 28 March 2024 as a Memorandum for the Heads of Executive Departments and Agencies *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. See: The White House, *Fact Sheet: Vice President Harris Announces OMB Policy to Advance Governance, Innovation, and Risk Management in Federal Agencies' Use of Artificial Intelligence*, THE WHITE HOUSE (Mar. 28, 2024) *available at*: https://www.whitehouse.gov/briefing-room/statements-releases/2024/03/28/fact-sheet-vice-president-harris-announces-omb-policy-to-advance-governance-innovation-and-risk-management-in-federal-agencies-use-of-artificial-intelligence/ (last visited Sep. 08, 2024); The White House, *Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence* (Oct. 30, 2023) *available at*: https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/ (last visited Sep. 08, 2024); Executive Office of the President (Office of Management and Budget), Memorandum for the Heads of Executive Departments and Agencies' (Mar. 28, 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.

[120] *Supra* note 119, Executive Office of the President (Office of Management and Budget), *Memorandum for the Heads of Executive Departments and Agencies*, Appendix I, point 2 ('(…)

Contd…

policing).[121] It is worth noting that, while this is a particularly promising policy binding federal agencies,[122] whether and the degree to which its provisions will be implemented at state level, where the actual policing often takes place, seems unclear.[123]

At state level, depending on the stage of the criminal procedure, where a technology may be used, there are rules imposing concrete obligations on law enforcement and criminal justice authorities. For the purposes of this contribution, it is worth referring to California's 2024 Rules of Court.[124] These Rules provide for adequate standard-setting on risk assessment technologies that are used for sentencing purposes.[125] The use of such risk assessments is relevant to the right to privacy, since

Purposes That Are Presumed to Be Rights-Impacting (…) b. In law enforcement contexts, producing risk assessments about individuals; predicting criminal recidivism; predicting criminal offenders; identifying criminal suspects or predicting perpetrators' identities; predicting victims of crime; forecasting crime; detecting gunshots; tracking personal vehicles over time in public spaces, including license plate readers; conducting biometric identification (…) sketching faces; reconstructing faces based on genetic information; monitoring social media; monitoring prisons; forensically analyzing criminal evidence; conducting forensic genetics; conducting cyber intrusions in the course of an investigation; conducting physical location-monitoring or tracking of individuals; or making determinations related to sentencing, parole, supervised release, probation, bail, pretrial release, or pretrial detention (…)').

[121] Executive Office of the President (Office of Management and Budget), 'Memorandum for the Heads of Executive Departments and Agencies' (n 119), section 5 (Managing risks from the use of artificial intelligence), under c ('(…) iv. Minimum Practices for Either Safety-Impacting or Rights-Impacting AI (…) agencies must follow these practices before using new or existing covered safety-impacting or rights-impacting AI (…) A. Complete an AI impact assessment. Agencies should update their impact assessments periodically and leverage them throughout the AI's lifecycle. In their impact assessments, agencies must document at least the following (…) 1. The intended purpose for the AI and its expected benefit, supported by specific metrics or qualitative analysis (…) 2. The potential risks of using AI, as well as what, if any, additional mitigation measures, beyond these minimum practices, the agency will take to help reduce these risks (…) 3. The quality and appropriateness of the relevant data (…)').

[122] *Supra* note 19, 2-4 (under 'a. Covered Agencies', 'b. Covered AI' and 'c. Applicability to National Security Systems').

[123] A discussion on pros and cons of the Policy in: Policing Project, *What Does the New White House Policy on AI Mean for Law Enforcement? Here Are Our Takeaways* (Apr. 16, 2024) *available at*: https://www.policingproject.org/news-main/2024/4/15/what-does-the-new-white-house-policy-on-ai-mean-for-law-enforcement-here-are-our-takeaways (last visited Sep. 08, 2024).

[124] *Supra* note 59, California Rules of Court, 'Standard 4.35. Court use of risk/needs assessments at sentencing.

[125] It is reminded that risk assessment technologies may not always use AI, but they can be

Contd…

it involves the processing of personal information; namely, static and dynamic risk factors (such as education, economic situation, gender, age, family or employment status) that are believed to enhance accuracy in predicting the risk of reoffending.[126] With the objective of minimising bias and the risk of recidivism, as well as enhancing public safety,[127] California's Rules of Court offer specific definitions (including the inputs and outputs of the technology);[128] require validation of risk assessment tools;[129] demand that risk assessments be solely considered in combination with other materials that independently support the sentencing decision;[130] impose on judges the duty to consider experts' comments on limitations;[131] guide judges on how to interpret risk assessments and their outputs;[132] and require adequate training.[133]

To conclude, in the context of law enforcement and criminal justice, the need for access to personal data by state authorities to effectively fight against crime and enhance public safety could to some extent justify limitations to the right to privacy. Still, given opacity engulfing AI-driven operations in these areas,[134] it is of utmost importance to have *ex-ante* and brightline rules regulating the processing of personal data. This would not only protect the right to privacy against unfair limitations (*e.g.*, unauthorised sharing or insufficient supervision), but also to enhance transparency, reliability and accuracy of AI-guided data processing.[135]

combined with AI in various ways with the objective of improving their performance (see clarifications and references in footnote 6).

[126] A long list of risk assessment technologies and relevant risk criteria they consider in: G Zara & D Farrington, CRIMINAL RECIDIVISM: EXPLANATION, PREDICTION AND PREVENTION 166 (Routledge 2016).

[127] *Supra* note 59, California Rules of Court, 'Standard 4.35. Court use of risk/needs assessments at sentencing point (a).

[128] *Supra* note 59, California Rules of Court, 'Standard 4.35. Court use of risk/needs assessments at sentencing, point (b).

[129] *Supra* note 59, California Rules of Court, 'Standard 4.35. Court use of risk/needs assessments at sentencing, points (b) and (c).

[130] *Supra* note 59, California Rules of Court, 'Standard 4.35. Court use of risk/needs assessments at sentencing, point (d).

[131] *Supra* note 59, California Rules of Court, 'Standard 4.35. Court use of risk/needs assessments at sentencing, point (d).

[132] *Supra* note 59, California Rules of Court, 'Standard 4.35. Court use of risk/needs assessments at sentencing, points (e) and (f).

[133] *Supra* note 59, California Rules of Court, 'Standard 4.35. Court use of risk/needs assessments at sentencing, point (g).

[134] *Supra* note 118, Hert & Bouchagair.

[135] As mentioned in the introduction of this section, law enforcement and criminal justice

Contd…

IX

What can the Global Observer Take Home from this US Analysis?

The American legal system is rich and complete, at least on paper. The constitutional basis and the regulatory framework 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!). In this paper, we discussed the benefits of an adversary system for challenging AI use in the court, the Federal Rules on Evidence (as amended in 2023), the development towards better judicial gatekeeping with regard to AI evidence in court; and this, in the spirit of the legacy of the US criminal justice system (the period of obliging acceptance, despite strict rules on evidence, seems to end). Our analysis included recent case law (e.g., *Puloka* or *Arteaga*), as well as regulatory interventions *(e.g.*, the 2023 amendments to the Federal Rule of Evidence 702, the 2024 AI Policy or California's 2024 Rules of Court).

Every use of AI in the criminal justice system is expected to be compliant with evidence-related law, as well as key constitutional rights, particularly the right to due process, equal protection, and privacy. The American legal landscape seems to fulfil this requirement but on paper only. Our discussion demonstrated that legal practice is far from perfect and that challenges remain. We quoted authorities in the field, pointing out that testing expert evidence in the court in the context of AI is still not perfect, especially in the absence of consistent forensic standards, external validation or industry standards. We also discussed authors, questioning ready acceptance of black box AI by law enforcement and calling for judicial rulings and legislation to safeguard a right to interpretable forensic AI and stop using black box AI.

Sabine Gless is in general positive about readiness of the US criminal procedure.[136] More than Europe, the US legal system and its case law have explored the question of what constitutes scientific knowledge and when an expert can be trusted in court. In Gless' terms, inquisitorial systems tend to blur what she calls 'the science dilemma'. Europeans know what science is when they see it, but Americans

authorities must have access to various items of information, including personal data, to more accurately investigate, adjudicate on or otherwise resolve a case. In the concrete context of artificial intelligence and its use in criminal procedure, the right to privacy and the protection of personal data becomes relevant, given that AI-related tools can process personal data in a rather sophisticated fashion; and there seems to be an acute need for bright line regulations on, *inter alia*, data integrity, quality, reliability, security, storage, retention or sharing. For a general discussion on privacy and open trials, see: Daniel Marshall and Terry Thomas, PRIVACY AND CRIMINAL JUSTICE 153 (Palgrave Macmillan 2017).

[136] *Supra* note 76

approach this more thoughtfully and can fall back on an elaborate set of rules and interpretations. On the negative side, adversarial systems, Gless observes, disregard the cost problem: one needs resources to bring in science and experts, resources that many Americans do not have. To conclude, some key take-homes can be summarised as follows:

Human experts to understand AI

AI usage that are neither comprehensible nor scrutinizable by human experts should, in our view, be avoided, when making critical decisions. This is something required by evidence-related law, demanding the involvement of an adequately qualified human expert, who can assist the court in understanding evidence or determining a matter, as well as by due process's general requirement to give reasons for decisions that should, moreover, be contestable. It would be desirable to avoid the use of AI, where decision-making procedures (*e.g.,* on guilt or sentencing) demand adequate and concrete reasons and humanly understandable explications that today's AI may fail to deliver.[137] Furthermore, regulators could be open and permissive towards AI-implementations in less critical contexts, where statistical precision is required and where no human rights are at stake (*e.g.,* uses by law enforcement to merely improve resolution of a low-quality video-footage). In such less critical cases, however, the goal pursued by a given AI-use should be limited to corroboration of a humanmade decision that should, moreover, be independently supported by other materials;[138] and AI's overall performance should be subjected to enhanced checks and balances, ensuring that transparency and other crucial safeguards are in place.[139]

Transparent AI

Transparency is imperative for the defence to effectively contest allegations against it and challenge their reliability, under the general requirements of due process. It

[137] Of relevance to this is explainability, requiring that AI-taken decisions be intelligible to humans. The extent to which adequate levels of explainability could be established remains unclear, given that, as any machine learning expert could confirm, some sophisticated forms of AI and their decisions can in no way be understood by humans.

[138] 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 (2016) 98.

[139] 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.

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; maximise accessibility, in case of IP-protected AI, via procedural legal instruments (like protective orders) granting conditional access without disproportionately limiting IP rights protection; 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.

Reliable AI

Some minimum standard-setting could ensure high levelled-reliability, *e.g.*, 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). The use of such tools could eliminate or reduce bias and error rates, enhance probative value of AI and make it more relevant, probably not outweighed by potential unfair prejudice or misleading. Such standard-setting could include permissible factors and inputs that could, to the relevant scientific community, promote accuracy; and exclude non-permissible factors/inputs that are, to the expert community, believed to result in unfair discrimination.[140]

Strict scrutiny testing for AI

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.

A legal framework for use of AI

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. The above-analysed legal instruments at the US federal and state level, endeavouring to impose concrete obligations with a view to protecting those who might become vulnerable to AI-

[140] 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).

uses,[141] demonstrate that detailed and clear lawmaking can promote the people's general interest in knowing and foreseeing state actions and their limits. The key lesson learnt from the US-privacy-section would probably suggest that regulators introduce brightline rules governing AI in an *ex-ante* fashion – before its actual application to the real world.[142] No 'hurt first, fix later' approach, as is currently the case in the U.K. and the rest of Europe. [143]

An *ex-ante* and brightline rulemaking approach need to be broadened to regulate, not only, issues directly connected to AI, but also to deal with issues impacted by AI. So, all issues in the four previous take-homes (experts' understanding and scrutiny, transparency, reliability and unfair-discrimination-strict-scrutiny-standard) would ideally be clarified and updated with a view to adequately protecting those who could be affected by AI-uses, before these people get actually affected.

[141] This probably includes a rather large number of people, given possible intrusiveness of contemporary AI-implementations. On vulnerability, see, in more detail: G Malgieri, VULNERABILITY AND DATA PROTECTION LAW (Oxford University Press 2023).

[142] Of relevance to this in Europe are the legal impact assessments provided for under the 2024 AI Act, the 2016 General Data Protection Regulation, as well as the 2016 Law Enforcement Directive. On the one hand, the AI Act's Fundamental Rights Impact Assessment (FRIA, in art 27 of the AI Act) sets out a rigorous assessment to be conducted for high-risk AI by (among others) public entities, as well as private entities providing public services. Still, it is doubtful whether it covers all AI-uses that could be included in the areas of law enforcement and criminal justice, given the lack of concrete definitions and the AI Act's indicative enlisting of private entities providing public services (see AI Act, recital 96). On the other hand, the Data Protection Impact Assessment (DPIAs, under the General Data Protection Regulation, art 35, and the Law Enforcement Directive, art 27) are not AI-targeted and are not always compulsory. It is added that, as Cohen has pointed out, the practical application of similar assessments may sometimes result in 'managerialisation' of privacy law and the reduction of its substantive goals to symbolic box-checking exercises. See: J Cohen, BETWEEN TRUTH AND POWER: THE LEGAL CONSTRUCTIONS OF INFORMATIONAL CAPITALISM 143-147 (Oxford University Press 2019). See also: A E Waldman, *Privacy Law's False Promise* 97(3) WASHINGTON UNIVERSITY LAW REVIEW 773 (2020) 776.

[143] M Leslie, J Summers, I Agerbak, ''Hurt first, fix later': AI regulation white paper consultation response' (2023) *available at*: https://publiclawproject.org.uk/resources/hurt-first-fix-later-ai-regulation-white-paper-consultation-response/ (last visited Sep. 08, 2024).