ARTICLES

AI IN THE COURTROOM: A COMPARATIVE ANALYSIS OF MACHINE EVIDENCE IN CRIMINAL TRIALS

Sabine Gless[*]

Abstract

As artificial intelligence (AI) has become more commonplace, the monitoring of human behavior by machines and software bots has created so-called machine evidence. This new type of evidence poses procedural challenges in criminal justice systems across the world due to the fact that they have traditionally been tailored for human testimony. This article's focus is on information proffered as evidence in criminal trials which has been generated by AI-driven systems that observe and evaluate the behavior of human users to predict future behavior in an attempt to enhance safety.

A poignant example of this type of evidence stemming from data generated by a consumer product is automated driving, where driving assistants as safety features, observe and evaluate a driver's ability to retake control of a vehicle where necessary. In Europe, for instance, new intelligent devices, including drowsiness detection and distraction warning systems, will become mandatory in new cars beginning in 2022. In the event that human-machine interactions cause harm (e.g., an accident involving an automated vehicle), there is likely to be a plethora of machine evidence, or data generated by AI-driven systems, potentially available for use in a criminal trial.

It is not yet clear if and how this the data can be used as evidence in criminal fact-finding, and adversarial and inquisitorial systems approach this issue very differently. Adversarial proceedings have the advantage of partisan vetting, which gives both sides the opportunity to challenge

[*] Sabine Gless is a Professor of Criminal Law and Procedure at the University of Basel School of Law (Switzerland) where she holds a Chair in Criminal Law and Criminal Procedure. She may be reached at Sabine.Gless@unibas.ch. This Article is the result of her participation in NYU's Hauser Global Scholarship Program during the Spring of 2019. Special thanks goes to all NYU School of Law faculty and staff for their support and the Swiss National Research Foundation for their ongoing support and funding, in particular the NFP75 Big Data grant. Furthermore, the author wishes to thank Gráinne de Búrca, Sara Beale, Eric Hilgendorf, Suzanne Kim, Erin Murphy, Richard Myers, Catherine Sharkey, Kathrin Strandburg, Thomas Weigend, Sarah Wood, and the participants of the Emile Noëlle & Hauser Workshop in March 2019 as well as the participants of the Data, Technology and Criminal Law Workshop at Duke Law School in April 2019 for their comments. © 2020, Sabine Gless.

consumer products offered as witnesses. By contrast, inquisitorial systems have specific mechanisms in place to introduce expert evidence recorded outside the courtroom, including to establish facts, which will be necessary to thoroughly test AI.

Using the German and the U.S. federal systems as examples, this Article highlights the challenges posed by machine evidence in criminal proceedings. The primary area of comparison is the maintenance of trust in fact-finding as the law evolves to accommodate the use of machine evidence. This comparative perspective illustrates the enigma of AI in the courtroom and foreshadows what will become inevitable problems in the not-too-distant future. The Article concludes that, at present, criminal justice systems are not sufficiently equipped to deal with the novel and varied types of information generated by embedded AI in consumer products. It is suggested that we merge the adversarial system's tools for bipartisan vetting of evidence with the inquisitorial system's inclusion of out-of-court statements under specific conditions to establish adequate means of testing machine evidence.

I. INTRODUCTION

Automated systems capable of handling a particular task, like driving a car, are currently defined as *narrow Artificial Intelligence (AI)*. This should be distinguished from general AI that possesses human-like cognitive abilities and an experiential understanding of its environments, coupled with the ability to process larger quantities of information at much greater speeds than the human mind.[1] This Article focuses on AI-driven systems that observe and evaluate the behavior of human users in order to predict future behavior, such as safety enhancing driving systems that react automatically and autonomously to the actions and reactions of human drivers, i.e. external information. The potential to use data generated by general AI technology in courtrooms poses novel challenges to both substantive criminal law and criminal procedure.[2] AI has the capacity to observe and assess humans' fitness to contribute to a wide range of cooperative actions. Will this result in another digital evidentiary gathering tool? Is such data sufficiently reliable to be used in criminal proceedings? Could such observations amount to a type of "machine testimony"[3] in the event of an accident? To address these questions, one must first acknowledge that robots and software bots—i.e., standalone machines or programs that interact with

[1] For a detailed discussion on definitional problems around AI, see Matthew U. Scherer, *Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies*, 29 HARV. J.L. & TECH. 353, 358–69 (2016).

[2] *See generally* Mireille Hildebrandt, *Ambient Intelligence, Criminal Liability and Democracy*, 2 CRIM. L. & PHIL. 163 (2008) (discussing the impact of ambient intelligence on criminal law).

[3] Andrea Roth, *Machine Testimony*, 126 YALE L.J. 1972, 1979 (2017).

users of a consumer product—are different from forensic instruments like breathalyzers, DNA testing kits, or radar speed guns. While the latter were designed to measure specific input data and perform straightforward calculations or provide other (predictable) output, narrow AI embedded in consumer products has the ability to collect information from a wide variety of inputs, assess the information autonomously for patterns, and convey a message based on algorithms and machine learning that is neither guided by nor entirely comprehended by humans. They were also not designed for evidentiary purposes. Thus, this message may be difficult to categorize and analyze using traditional evidentiary rules.

An analysis of the relevant law on this topic reveals that the fact-finding procedure, and particularly the assessment of evidentiary reliability, is a human-focused phenomenon with the goal of providing transparent and objective information to the trier of fact while also safeguarding a reliable and valid fact-finding process. Therefore, the use of narrow AI in forensic instruments already poses challenges to evidentiary law and the appraisal of fact today. For instance, digitized breathalyzers have shed light on issues surrounding these types of evidentiary assistants that contain inherent black-box problems—an inability to adequately explain their inner workings.[4]

If the data generated by AI during a collaborative act with a human is admissible in court, it could potentially be classified as a form of documentary evidence or even a type of witness testimony. Regardless, the underlying issue remains: how AI and the data it produces can be meaningfully evaluated for reliability and credibility, particularly when the data presented in court has been generated by technology that evaluates human behavior, not in an effort to produce tangible evidence, but rather to meet a specific commercial need without taking into consideration issues surrounding criminal justice systems.

An investigation of this issue in the German and the U.S. federal systems will illustrate that scrutiny in fact-finding is much more complex in an adversarial system, where there are a plethora of ways to test the credibility and reliability of evidence, and where scholars have already

[4] For more details on digitized breathalyzers, see *id.* at 1972, 1979, 2015–16; Stacy Cowley & Jessica Silver-Greenberg, *These Machines Can Put You in Jail. Don't Trust Them*, N.Y. TIMES (Nov. 3, 2019), https://www.nytimes.com/2019/11/03/business/drunk-driving-breathalyzer.html.

suggested solutions for new generations of digital evidence.[5] Because of an inherent lack of means to evaluate reliability and credibility of evidence in an inquisitorial system like that of Germany, a new legal pathway to do so would need to be established should it endeavor to create means of systematically evaluating machine evidence in criminal fact-finding. That said, the inquisitorial tradition of gradually building a case file and relying on out-of-court statements could more readily allow a thorough evaluation and simultaneous paper documentation of complex evidence testing that would be available to all parties from the beginning. Examining the performance of an AI-driven tool requires time and successive trial runs, in addition to experimenting with the machine and a detailed record showing all results. A thorough examination by a court-appointed expert who provides the results to both parties is potentially a more feasible way to evaluate evidence that, despite being technically complex, must still be presented and explained orally in a courtroom.

This Article argues, from a comparative legal angle, that neither the inquisitorial systems prevalent on the European continent, nor the adversarial system used in the United States, are prepared for AI in the courtroom and thus cannot take advantage of potentially relevant machine evidence. While inquisitorial systems have struggled to find adequate defense tools to combat this new form of information, adversarial systems have few feasible means of including out-of-court tests documenting a thorough vetting of AI-driven devices. This Article proposes significant changes to both systems in anticipation of courts across the world being faced with evidence generated by AI and argues for an approach that draws from both adversarial and inquisitive legal systems. This would include the adversarial systems' tools to thoroughly scrutinize evidence in a partisan setting and the inquisitorial systems' allotment of the time and space needed to assess complex technical evidence outside the courtroom, and its sharing of knowledge among all parties in a case file. The ultimate goal of the Article is to provide a new approach for the presentation of machine evidence in a criminal trial.

The Article first provides a brief sketch of machine evidence generated by AI, using the primary example of traffic accidents involving automated driving, whereby data monitoring a human driver's face for drowsiness, the activation of a drowsiness alert, or the driving assistant's

[5] *See* Erin Murphy, *The New Forensics: Criminal Justice, False Certainty, and the Second Generation of Scientific Evidence*, 95 CALIF. L. REV. 721 (2007); Brandon L. Garrett, *Big Data and Due Process*, 99 CORNELL L. REV. ONLINE 207 (2014); Christian Chessman, *A "Source" of Error: Computer Code, Criminal Defendants, and the Constitution*, 105 CALIF. L. REV. 179 (2017).

assessment of a driver's conduct could all be potentially relevant evidence. Second, it analyzes how machine evidence generated by AI could be introduced into trial by way of experts reporting on their findings or by devising a way to bring the software or machine into the courtroom. From the point of ensuring trustworthiness in fact-finding, it is difficult to simply introduce machine evidence as some form of documentary evidence or relate it to testimonial evidence. Machine evidence may not fall in either category, but instead demands a new evidentiary approach that takes into account that AI, like human witnesses, could identify a particular defendant as the perpetrator of the crime based on its own evaluation. Therefore, it must be vetted as a witness rather than as a tool providing a test result.

A. *Legally Defining AI*

Legal scholars are not the only ones that continue to struggle to define technological terms like AI,[6] robot,[7] or bot.[8] This difficulty points to the rapid and significant developments in technology and suggests that the law has not yet caught up. However, the lack of a statutory definition for AI should not inhibit an analysis of how ambient intelligent environments, i.e. those where electronic devices are capable of monitoring and responding to human behavior, might impact criminal justice. Quite the contrary, an agreed-upon legal definition will only be possible if attorneys and legal scholars discuss the various aspects of new technology and its potential impact on legal systems. For the purpose of this Article, the focus is on AI-driven systems that observe and evaluate the behavior of human users to predict future behavior in an attempt to enhance safety while reacting automatically and

[6] J. McCarthy et al., *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence* (Aug. 31, 1955), http://jmc.stanford.edu/articles/dartmouth/dartmouth.pdf. For a more recent discussion, see Yavar Bathaee, *The Artificial Blackbox and the Failure of Intent and Causation*, 31 HARV. J.L. & TECH. 889, 898 (2018).

[7] *See* Matt Simon, *What is a Robot?*, WIRED (Aug. 24, 2017), https://www.wired.com/story/what-is-a-robot/ (describing a range of disagreement); *see also* Ryan Calo, *Robots as Legal Metaphors*, 30 HARV. J.L. & TECH. 209, 215 (2016).

[8] For purposes of this Article, bots will be defined as automated software capable of interacting with human users or other IT systems. *See* Carlene R. Lebeuf, A Taxonomy of Software Bots: Towards a Deeper Understanding of Software Bot Characteristics (Aug. 31, 2018) (unpublished M.S. thesis, University of Victoria) (on file with the University of Victoria Libraries) (discussing, thoroughly, various definitions of software bot and proposing a new taxonomy); *see, e.g.*, Renee DiResta, *A New Law Makes Bots Identify Themselves -That's the Problem*, WIRED (July 24, 2019), https://www.wired.com/story/law-makes-bots-identify-themselves (highlighting the potential consequences of California lawmakers' definition of 'bot').

autonomously.[9] Such systems can take the shape of a robot or software bot,[10] but always possess their own agenda, automaticity, and ability to autonomously evaluate.[11]

B. *Methodology*

This Article uses two approaches. The first is an analysis of the substantive and procedural law around machine evidence in criminal trials and the issues that will likely arise regardless of jurisdiction or type of legal system. The second is a functional comparative approach based on the original work of Zweigert & Kötz.[12] The *tertium comparationis*, or most relevant point of comparison, is the means by which trustworthiness is evaluated during criminal fact-finding where machine evidence is presented. The German and U.S. federal systems will serve as examples as they represent contrasting criminal justice systems and both are relevant as car-manufacturing nations.

This Article analyzes statutory provisions, evidentiary rules, and relevant case law pertaining to trustworthiness in fact-finding in each of the two jurisdictions, particularly where automated machine-generated evidence is at issue. The use of experts and technical reports is also examined and illustrates opposing approaches to such evidence. While such a legal comparison can never be completely neutral because the meaning of any one term can vary vastly across cultures and jurisdictions, it does allow for the incorporation of specific judicial concepts within an overarching legal reality.[13] Here, the concept of trustworthy fact-finding is closely tied to divergent underlying values (i.e., trust in citizen jurors and judges in the United States versus benches exclusively comprised of judges in Germany) and is laden with normative layers that may distort the validity of the comparative findings if not contextualized properly.[14]

[9] Mark A. Lemley & Bryan Casey, *You Might Be a Robot*, CORNELL L. REV. (2019), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3327602; Chessman, *supra* note 5, at 206, 220 n.310; *see* Esra Vural et al., *Drowsy Driver Detection Through Facial Movement Analysis*, *in* HUMAN-COMPUTER INTERACTION 6–18 (Michael Lew et al. eds., 2007) (describing drowsiness detection systems).

[10] Ryan Calo, *Robotics and the Lessons of Cyberlaw*, 103 CALIF. L. REV. 513, 530–31 (2015).

[11] *Id.*

[12] KONRAD ZWEIGERT & HEIN KÖTZ, AN INTRODUCTION TO COMPARATIVE LAW 33 (3d ed. 1998).

[13] Axel Tschentscher, *Dialektische Rechtsvergleichung–Zur Methode der Komparatistik im öffentlichen Recht*, 17 JURISTENZEITUNG 807 (2007) (Ger.).

[14] Annelise Riles, *Wigmore's Treasure Box: Comparative Law in the Era of Information*, 40 HARV. INT'L L.J. 221, 225 (1999); For an example of comparative contextualization, see Mirjan R. Damaška, *Of Hearsay and Its Analogues*, 76 MINN. L. REV. 425, 431 (1991–1992).

It is the objective of any functional approach[15] to look beyond the normative layers and serve as a tool to analyze the specific factual problems of machine evidence in criminal courts, referencing the process tied directly to the more ubiquitous issue of how to evaluate machine evidence, which will eventually be an issue for all jurisdictions.

II. Machine Evidence Generated by AI

The following paragraphs argue that automated driving raises novel, yet intertwined, issues in a variety of criminal justice domains. Unresolved issues of criminal responsibility in substantive law will lead to evidentiary problems where human-robot interactions cause harm because a lack of clarity around whether AI shares criminal responsibility with the human driver results in confusion about how AI-generated machine data shall be treated. Is it to be considered a witness providing expert or percipient testimony or should it be categorized as a statement by a co-defendant?

A. *Automated Driving*

Automated driving is an everyday example of the development of technology that has led to AI monitoring humans. As this technology progresses, humans will increasingly be sharing the wheel with so-called "driving assistants," or software bots that support the human driver's performance and assist or even take over driving in specific situations.[16] In the case of the latter, it is unclear, however, who will be seen as the driver at any given moment, and this has significant consequences for liability.[17]

Automated driving technology is already capable of carrying out complex series of actions independent of the human driver. These systems monitor the vehicle's position in the lane and the driver's steering pattern, body temperature, and facial movements (particularly ocular

[15] Oliver Brand, *Conceptual Comparisons: Towards a Coherent Methodology of Comparative Legal Studies*, 32 BROOK. J. INT'L L. 405, 415 (2007) (promulgating another functional comparative theory, a variant of which this author uses).

[16] *See, e.g.*, Markus Enzweiler, *The Mobile Revolution–Machine Intelligence for Autonomous Vehicles*, 57 INFO. TECH. 199 (2015).

[17] Sabine Gless, Emily Silverman, & Thomas Weigend, *If Robots Cause Harm, Who is to Blame? Self-driving Cars and Criminal Liability*, 19 NEW CRIM. L. REV. 412 (2016); Susanne Beck, *Robotics and Criminal Law. Negligence, Diffusion of Liability and Electronic Personhood*, *in* DIGITIZATION AND THE LAW 41, 46 (Eric Hilgendorf & Jochen Feldle eds., 2018); Sabine Gless & Thomas Weigend, *Intelligente Agenten und das Strafrecht*, 126 ZEITSCHRIFT FÜR DIE GESAMTE STRAFRECHTSWISSENSCHAFT 561, 578 (2014) (Ger.).

movement).[18] They are able to learn a driver's typical posture, head position, blink rate, facial expressions, and steering patterns. Where anomalies are detected, the driver is warned to stop and take a break (e.g., with a blinking orange coffee cup sign). While some driver assistance technology may appear almost toy-like, systems that monitor a human's driving behavior are crucial where the human relinquishes control of the vehicle. At this stage, the automated driving technology is the primary driver, but the human nevertheless must respond to a request to intervene and take over control if the driving assistants cannot handle a particular situation.[19] Therefore, take-over-request (TOR) devices are constantly monitoring whether the human driver appears capable of doing so when necessary.[20]

As many accidents are caused by sleepy drivers, drowsiness detection systems may be considered crucial safety features capable of observing human drivers and recording their actions and reactions. The EU revised its General Safety Regulations to designate drowsiness detections systems as mandatory safety features in European vehicles beginning in 2022.[21] The technology is likely to build upon the rapid development and impressive success of facial recognition technology and is part of a growing industry investing in machine-human interfaces involving human monitoring.[22]

Although futuristic, cars with standard drowsiness detention systems and other features able to continually monitor human drivers are not science fiction. In addition to the EU, a number of jurisdictions already

[18] Yanchao Dong et al., *Driver Inattention Monitoring System for Intelligent Vehicles: A Review*, 12 IEEE TRANSACTIONS ON INTELLIGENT TRANSP. SYS. 596 (2011); Luis M. Bergasa et al., *Real-Time System for Monitoring Driver Vigilance*, 7 IEEE TRANSACTIONS ON INTELLIGENT TRANSP. SYS. 63 (2006).

[19] Madeline Roe, Who's Driving That Car? An Analysis of Regulatory and Potential Liability Frameworks for Driverless Cars, 60 B.C.L. REV. 315, 319 (2019).

[20] *Cf.* Vivien Melcher et al., *Take-Over Requests for Automated Driving*, 3 PROCEDIA MANUFACTURING 2867 (2015); Tobias Vogelpohl et al., *Asleep at the Automated Wheel—Sleepiness and Fatigue During Highly Automated Driving*, 126 ACCIDENT ANALYSIS & PREVENTION 70 (2018); Joel Gonçalves et al., *Drowsiness in Conditional Automation: Proneness, Diagnosis and Driving Performance Effects*, 2016 INST. OF ELECTRICAL AND ELECTRONICS ENG'RS (IEEE) 19TH INT'L CONF. ON INTELLIGENT TRANSP. SYS. (ITSC) 873.

[21] *See* Regulation (EU) 2019/2144 of 27 November 2019 on Type-Approval Requirements for Motor Vehicles and their Trailers, and Systems, Components and Separate Technical Units Intended for such Vehicles, as Regards their Safety and the Protection of Vehicle Occupants and Vulnerable Road Users, 2019 O.J. (L 325) 1.

[22] *See* Jasper Gielen & Jean-Marie Aerts, *Feature Extraction and Evaluation for Driver Drowsiness Detection Based on Thermoregulation*, APPLIED SCIENCES 2019, Aug. 30, 2019, at 3555; for information about EU technology, see also EUROPEAN COMMISSION, IN-VEHICLE DETECTION AND WARNING DEVICES, https://ec.europa.eu/transport/road_safety/specialist/knowledge/fatique/countermeasures/in_vehicle_detection_and_warning_devices_en (last visited Feb. 6, 2020).

allow for human-robot interaction in automated driving. For instance, in June 2017, Germany passed a regulation on this issue,[23] and the Swiss Administration is currently considering laws that eventually allow for fully automated vehicles.[24] Similar laws are also found in several U.S. states.[25]

When is a device simply a tool and when does it reach the level of a software bot or robot? Driving automation and the use of driving assistants exemplifies both a dividing line and the gray area around the use of AI for evidentiary purposes. Using a pragmatic functionality test that looks at the factual problems in using such evidence,[26] one can differentiate between an AI-driven device that does not convey any message of its own (i.e., one that is solely a tool serving a human user) and when such a degree of autonomy exists that the information produced would be considered AI-generated machine evidence rendering an opinion (for instance, assessing a human's capability to drive a car). Only the latter would require separate credibility testing in the courtroom.[27] Since the 1950s, anti-lock brake systems (ABS) have been routinely used in automobiles. This safety feature engages automated stutter braking designed to prevent the wheels from locking and to maintain contact with the road more effectively than with the driver's braking alone.[28] Anti-lock brakes have become increasingly sophisticated but the technology is still seen as something that merely responds to a non-human entity (the road conditions) and is, therefore, a tool to be used by a driver that does not add a message of its own, i.e., it does not evaluate human performance during the braking action and provide an opinion.

[23] Achtes Gesetz zur Änderung des Straßenverkehrsgesetzes [Eighth Amendment to the Road Traffic Law], BGBL I at 1648 (Ger.).

[24] *See* Aktivitäten des Bundes [Federal Activities], BUNDESAMT FÜR STRASSEN (ASTRA), https://www.astra.admin.ch/astra/de/home/themen/intelligente-mobilitaet/aktivitaeten-des-bundes-.html (last accessed Mar. 24, 2020).

[25] For statistics and an autonomous vehicles legislation database, see NATIONAL CONFERENCE OF STATE LEGISLATURES, AUTONOMOUS VEHICLES: SELF-DRIVING VEHICLES ENACTED LEGISLATION, www.ncsl.org/research/transportation/autonomous-vehicles-self-driving-vehicles-enacted-legislation.aspx (last visited Feb. 6, 2020).

[26] Lemley & Casey, *supra* note 9.

[27] AI-driven devices, however, differ on the level of conveying a claim of their own. Thus, it will be important to differentiate among their complexity, opacity, sensitivity to (case-specific) human manipulation, and the concrete use in a case. *See, e.g.*, Roth, *supra* note 3, at 1979, 1986–93, 2002–22.

[28] The automation technology operates at a much faster rate and more effectively than most drivers could manage.

In comparison, drowsiness detection systems are driving assistant software bots designed to enhance driving safety by observing and assessing a human driver's behavior and alerting the driver when he or she appears drowsy. These systems can be used when the driver maintains control of the vehicle or they may be part of a TOR assistant. The technology at work may include surveillance of steering patterns, lane position, facial changes in the driver, and the driver's sitting position.[29] In contrast to ABS, drowsiness detection systems add their own message; that is, if they evaluate a driver to be sleepy they will issue an alert. This alert is then recorded and stored and could potentially be used as evidence in a criminal trial.

B. *Substantive Law*

In recent years, smart devices operating through the use of AI, machine learning, and big data, have begun to create new opportunities in many personal and professional domains. Today, digital assistants can help medical doctors detect patterns indicative of certain illnesses; smart houses can provide assistance to aging residents, individuals with disabilities, or anyone seeking more convenience; and automated driving can offer those wanting a little extra free-time during their commute the option to share the responsibility of driving.[30]

The AI technology that created these new possibilities took roughly 50 years to become public. When scholars coined the term AI in the 1950s, they were referring to the science and engineering behind the creation of "intelligent machines."[31] It is unlikely they would have anticipated the capacity of AI, which can learn to re-organize itself to improve efficiency, including rewriting its own code.[32] Nevertheless, the progress in AI has not been generalized to more holistic problem-solving strategies, and instead deals only with specific tasks, which is one of the reasons it is called *narrow AI*.[33]

Increases in new possibilities for human-robot interaction also increase the possibility of harm as a result of such cooperation, even

[29] *See, e.g.*, Chris Schwarz, John Gaspar, Thomas Miller & Reza Yousefian, *The Detection of Drowsiness Using a Driver Monitoring System*, 20 TRAFFIC INJ. PREVENTION 157-161 (2019).

[30] The available technology does not stop there—law enforcement agencies also make wide use of new tools, for example, to calculate risk when deciding about bail or release from prison and when pursuing suspects or potential offenders. Andrea Roth, *Trial by Machine*, 104 GEO. L.J. 1245 (2016).

[31] McCarthy, *supra* note 6; Bathaee, *supra* note 6.

[32] Calo, *supra* note 10, at 534.

[33] Scherer, *supra* note 1, at 358–69.

though the expectation is that automated driving or the use of driving assistants will enhance safety on the roads. Where such actions cause harm, it may become necessary to assign blame.[34] In that respect, it could be argued that AI-driven devices (or their creators) could be viewed as potential defendants that share responsibility with human users.

As things stand today, we do not consider any form of AI a moral agent able to stand trial. Its intelligence is largely one-dimensional and lacks, at least for the time being, the capacity to reflect and account for past actions.[35] Nevertheless, developments in technology have sparked a scholarly debate on robot liability that acknowledges while blame is a social construct, at some point it could include machines should society agree that robots are suitable agents of responsibility.[36] The likelihood of this development seems to be related to the degree to which AI is able to accurately assess information across a variety of domains, as well as their ability to develop some sort of reasoning akin to human common sense.

Even if one subscribes to the traditional views denying AI agency,[37] a lack of legal liability does not necessarily mean that it must be regarded as a neutral bystander. In some ways, a robot or software bot could be seen as a secondary (maybe not always legally responsible) suspect, or even a proxy suspect for those manufacturing the automobile. So, while we view AI as lacking agency and a moral compass, it can have faults that most would expect to trigger some form of liability. A drowsiness detection system, for instance, can be imprecise or ambiguous—it may include biased algorithms or standardized data, or something else entirely. Along this line, some AI has been shown to have an "automation bias" in software design favoring the corporate self-interest.[38]

Questions around who is responsible when someone is harmed by a car operating autonomously and how to allocate guilt are closely connected.[39] Challenges related to substantive law were the first legal issues

[34] Gless, Silverman & Weigend, *supra* note 17.

[35] Dafni Lima, *Could AI Agents Be Held Criminally Liable? Artificial Intelligence and the Challenges for Criminal Law*, 69 S.C.L. REV. 677 (2018); Ying Hu, *Robot Criminals*, 52 U. MICH. J.L. REFORM 487 (2019); *see also* John C. Coffee, *No Soul to Damn: No Body to Kick: An Unscandalized Inquiry into the Problem of Corporate Punishment*, 79 MICH. L. REV. 386 (1981) (analyzing, broadly, punishment of non-humans).

[36] Monika Simmler & Nora Markwalder, *Guilty Robots? Rethinking the Nature of Culpability and Legal Personhood in an Age of Artificial Intelligence*, 30 CRIM. L.F. 1 (2019).

[37] Gless, Silverman, & Weigend, *supra* note 17, at 412.

[38] *Id.*

[39] Scherer, *supra* note 1, at 358, 366–67; Wolfgang Wohlers, *Individualverkehr im 21. Jahrhundert: das Strafrecht vor neuen Herausforderungen*, 3 BASLER JURISTISCHE MITTEILUNGEN 113 (2016) (Ger.).

to emerge, and it is likely that our conceptualization of agency, how guilt should be allocated, and who constitutes a perpetrator or accomplice will have to change with more AI in our lives. Until now, substantive criminal law has tended to focus on humans as moral agents, capable actors, and occasionally criminal risk-takers. While it is true that domestic lawmakers have become more receptive to the idea that non-humans, such as corporations, can be criminally responsible, prosecutions continue to be rooted in the idea that only human action is subject to criminal liability.[40]

C. *AI and the Evidentiary Cycle in Criminal Trials*

Machine evidence, like other forms of technology that came before, has the potential to provide new sources of information and, thus, a chance for more accurate fact-finding in criminal trials. However, the use of technology with inherent black box problems, i.e., an inability to explain a certain result, in a criminal proceeding comes at a price. Triers of fact will have to decide whether to trust an AI-generated statement that can only partially be explained by experts. In the past, courts have opposed the use of forensic tools, like breathalyzers, noting that they act as "magic black box[es] assisting the prosecution in convicting citizens." [41] Courts would be wise to be skeptical as AI becomes more embedded in future generations of forensic tools. Machine evidence generated by AI in consumer products, such as driving assistants, poses new challenges in light of the fact that it was developed as a solution to a consumer need and was not meant to be used as a forensic evidentiary tool.

It may initially appear unlikely that increased AI use in our daily lives would result in the increased importance of machine evidence in criminal trials to establish facts, particularly given courts' hesitation to use all available technology in the past (e.g., polygraphs), but continued technological development may result in a shift in judges' attitudes. As AI becomes more ubiquitous, and if such technology is deemed to be an accurate assessment of human conduct, more people may be willing to accept it as a reliable and trustworthy source of information. Despite this possibility, it remains unclear if and how such information would be admitted into a court of law. Long before AI came along, other

[40] *See* Sabine Gless & Sylwia Broniszewska-Emdin (eds.), *Prosecuting Corporations for Violations of International Criminal Law: Jurisdictional Issues*, 88 INT'L REV. OF PENAL L. (SPECIAL ISSUE) 1 (2017) (providing a comparative overview).

[41] State v. Lance, No. 48-2012-CT-000017-A /A, slip op. at 24 (Fla. Orange County Ct. Sept. 22, 2014).

technological developments were creating new forms of evidence (e.g., DNA testing) that, once brought into criminal trials, highlighted the unpreparedness of the criminal procedural process.[42]

1. From Silent Witnesses to Digital Analytical Tools

With the emergence of every new type of recording technology, courts face the question of whether the means of registration or documentation is reliable, accurate, and objective.[43] Going forward, this process will be referred to as the evidentiary life cycle.[44] These cycles are of fundamental interest because the legal questions that arise go beyond the evidentiary level to important constitutional issues like the Confrontation Clause. They also raise basic questions with regard to the use of new technology for fact-finding in criminal trials and the foundation for expert evidence. From a comparative point of view, it is interesting to note that presently the debates in the United States appear to be more focused on the right to confront adverse testimony[45] while in Germany the argument is (still) predominantly framed as a privacy issue.[46]

[42] Murphy, *supra* note 5, at 728–44.

[43] *See* James E. Bibart, *Metadata in Digital Photography: The Need for Protection and Production of this Silent Witness*, 44 Cap. U.L. Rev. 789 (2016).

[44] *See infra* Part III.B.3.

[45] *See, e.g.*, Murphy, *supra* note 5, at 775; Roth, *supra* note 3, at 1979, 1986–93, 2002–22; Joseph C. Celentino, *Face-to-Face with Facial Recognition Evidence: Admissibility Under the Post-Crawford Confrontation Clause*, 114 Mich. L. Rev. 1317, 1331–33 (2016). For debates regarding privacy issues, *see* Katherine Strandbergh, *Home, Home on the Web: The Fourth Amendment and Technosocial Change*, 70 Md. L. Rev. 614 (2011); Jason M. Weinstein, William L. Drake & Nicholas P. Silverman, *Privacy vs. Public Safety: Prosecuting and Defending Criminal Cases in the Post-Snowden Era*, 52 Am. Crim. L. Rev. 729 (2015).

[46] Bundesverfassungsgericht [BVerfGE] [Federal Constitutional Court] Feb. 2, 2008, 1 BvR 370/07 (§ 169) (Ger.); Bundesverfassungsgericht [BVerfGE], [Federal Constitutional Court] Mar. 3, 2004, 1 BvR 2378/98 and 1 BvR 1084/99 (Ger.); Bundesverfassungsgericht [BVerfGE] Apr. 4, 2006, 1 BvR 518/02 (Ger.); *see generally* Jan-Christoph Wehage, Das Grundrecht auf Gewährleistung der Vertraulichkeit und Integrität von Informationsverarbeitungssystemen (2013) (Ger.); *see also* Tobias Singelnstein & Benjamin Derin, *Das Gesetz zur effektiveren und praxistauglicheren Ausgestaltung des Strafverfahrens*, 70 Neue Juristische Wochenschrift 2646, 2647–52 (2017) (Ger.); Sabine Gless, *Wenn das Haus mithört: Beweisverbote im digitalen Zeitalter*, 38 Strafverteidiger 671, 675–77 (2018) (Ger.). A more recent debate does however include considerations on reliability: Oberlandesgericht [OLG] [Higher Regional Court] Karlsruhe, July 16, 2019, 1 Rb10 Ss 291/19 (Baden Württemberg) (Ger.); Verfassungsgerichtshof [BayVerfGH] [Constitutional Court] Apr. 27, 2018, Lv 1/18 (Saarbrücken) (Ger.); Eric Hilgendorf,"*Die Schuld ist immer zweifellos*"*? Offene Fragen bei Tatsachenfeststellung und Beweis mit Hilfe ,,intelligenter" Maschinen, in* Beweis 229, 238–39 (Thomas Fischer ed., 2019) (Ger.); Dominik Brodowski, *Die Beweisführung mit digitalen Spuren und das*

It is hoped that eventually legal scholars will join in a single debate around the relevance and reliability of AI-generated data during interactions with humans as well as on the resulting privacy issues. On the one hand, some may argue in favor of this new and seemingly more precise assessment instrument, hoping for a more accurate establishment of facts and laying aside privacy and other concerns. On the other hand, critics might describe such technology as invasive and error-prone, citing flaws in its design and/or the machine learning technology used. It is the functionality of the technology that draws the line between classic silent witnesses (like an analog video camera or an ABS breaking system), digital forensic analytical tools (like DNA testing kits or forensic facial recognition), and AI-driven devices (like drowsiness detection systems), which convey messages of their own through an independent evaluation of the situation.

With technology of some form or another in all areas of our lives, it is not surprising that digital evidence has already made its way into criminal proceedings.[47] Interestingly, the term "digital evidence" is used widely in textbooks and legal journals, yet it does not appear to be a technical legal term but instead a description of a phenomenon (or need) in criminal investigations and courtrooms for information stored or transmitted in binary form.[48] Digital evidence can be found on hard drives of computers, in cloud-based storage, on mobile phones or personal digital assistants, flash drives, or even digital cameras. Digital evidence can provide access to a large variety of information, including the content of an email, the identities of senders and recipients of emails, surveillance reports from camera footage, mobile location tracking records,[49] browser tracking information, or social network mapping data. It can be "small data" or some aspect of big data[50] and it can be generated by humans or machines. Currently, it is most often used to prosecute crimes[51] but, like DNA testing, it could eventually

Unmittelbarkeitsprinzip, in DIGITALISIERUNG DER GERICHTLICHEN VERFAHREN UND DAS PROZESSRECHT 83–93 (Almuth Buschmann et al. eds., 2018) (Ger.) (discussing the application of best evidence rules in digital evidence in German procedure).

[47] Orin S. Kerr, *Digital Evidence and the New Criminal Procedure*, 105 COLUM. L. REV. 279, 285–306 (2005).

[48] *See, e.g.*, Donald A. Dripps, *"Dearest Property": Digital Evidence and the History of Private "Papers" as Special Objects of Search and Seizure*, 103 J. CRIM. L. & CRIMINOLOGY 49, 53–54 (2013); Kerr, *supra* note 47.

[49] *See, e.g.*, Joshua A.T. Fairfield & Erik Luna, *Digital Innocence*, 99 CORNELL L. REV. 981, 1032–38 (2014).

[50] Garrett, *supra* note 5, at 208–09.

[51] Historically, new technology touted as "objective" was first used against defendants, but more recently defendants have (at least in part) turned this concept around. Roth, *supra* note 30, at 1254–64.

become helpful for the defense as well,[52] as users are increasingly aware of the ability of consumer products to provide alibis.

Information coded in "0s" and "1s" poses the immediate problem for the court in that it has to be translated into an analogue form by experts to be of any use.[53] Sometimes the expert can only access the information after it has been decrypted, and additional specialists are needed to provide explanations about how it was obtained and what it means. Nevertheless, so long as the issue is solely the storage format—that is, the information is simply in a different form (e.g., an email instead of a written note, a jpg file instead of a printed photo, etc.), criminal justice systems have been able to adapt without too much difficulty.

However, technology has continued to develop rapidly, and digital evidence in criminal investigations quickly became more complex: breathalyzers were equipped with smart technology, DNA kits providing personalized genetic information were made available to the public, and law enforcement no longer had exclusive access to such technology.[54] Such digital analytical tools can be AI-driven but are first-and-foremost measuring tools and are limited in that they are used to produce test results such as the alcohol content in someone's breath.[55] Even so, the basic question remains: when (and why) does the trier of fact start to trust such technology? When do we get to a point where the technology becomes so trusted that it is no longer sufficiently challenged? And when could new doubts arise about its trustworthiness?

In 2007, Erin Murphy presented a taxonomy dividing first-generation forensic evidence (e.g., fingerprint analysis, ballistics) from second-generation evidence,[56] primarily based upon the first generation's limited application, observational and mechanical functions, and narrow design. Second-generation forensic evidence was characterized as more complex and scientifically robust, which resulted in much broader use (e.g., DNA-analysis, location tracking).

[52] *See* Fairfield & Luna, *supra* note 49, at 990 (championing the demand to turn digital devices into proof-of-innocence technologies).

[53] Kerr, *supra* note 47, at 298–99.

[54] *See, e.g.*, Stacy Cowley & Jessica Silver-Greenberg, *These Machines Can Put You in Jail. Don't Trust Them*, N.Y. TIMES (Nov. 3, 2019), https://www.nytimes.com/2019/11/03/business/drunk-driving-breathalyzer.html; Matthias Gafni & Lisa M. Krieger, *Here's the 'Open-Source' Genealogy DNA Website That Helped Crack the Golden State Killer Case*, MERCURY NEWS (Apr. 30, 2019), www.mercurynews.com/2018/04/26/ancestry-23andme-deny-assisting-law-enforcement-in-east-area-rapist-case (illustrating the successful use of an "open source" genealogy website).

[55] Roth, *supra* note 30, at 1254–64 (providing a history of such tools).

[56] Murphy, *supra* note 5.

Digital analytical tools in forensic settings can be distinguished from first-generation evidence (like dactyloscopy or graphology) because they are guided by source code rather than human expertise[57] and their underlying mechanisms are considerably more difficult to see at work.[58] This lack of transparency alone creates a substantial risk to trustworthiness in criminal fact-finding because potential flaws are difficult to detect by the trier of fact. While the trustworthiness of DNA tests has recently triggered a lively debate, the underlying technologies are quite different because DNA tests lack the agency that AI can achieve, including the ability to monitor the surrounding environment, evaluate human behavior and act autonomously. As such, machine-generated evidence must be considered a third-generation type of forensic evidence.

2. Digital Layers and Trustworthy Fact-Finding

With each additional layer of digital complexity, access to relevant information becomes more difficult and requires expertise that the trier of fact might not possess. Additional issues arise from laws regulating the reliability and credibility of evidence which impact not only the admissibility of evidence (a key tenant of adversarial proceedings), but also its weight (a particularly important aspect in establishing facts in the inquisitorial model).[59]

a. The Black Box Problem and Expert Evidence

While a specific component of a drowsiness detection system in a car might be visible to the user, the process behind its evaluation is not entirely transparent to the user or a trier of fact in a courtroom. Even experts called upon to explain machine evidence in court encounter limitations in their ability to comprehensibly explain how an AI-driven device evaluates a human user's conduct or demonstrate a clear chain of causality.[60] These problems constitute the "black box problem" in machine evidence that researchers are currently investigating.[61] To date, research has shown that the degree to which AI can be explained

[57] Source code is the fundamental component of the IT program driving the action created by a human programmer.

[58] Roth, *supra* note 30, at 1269–76.

[59] Thomas Weigend, *Evidence Law im anglo-amerikanischen Strafverfahren*, *in* BEWEIS 253–65 (Thomas Fischer ed., 2019) (Ger.).

[60] *See* Finale Doshi-Velez & Mason Kortz, *Accountability of AI Under the Law: The Role of Explanation* (Berkman Klein Ctr. Working Grp. on Explanation and the Law, Berkman Klein Ctr. for Internet & Soc'y, Working Paper, 2017).

[61] *See* Cary Coglianese & David Lehr, *Regulating by Robot: Administrative Decision Making in the Machine-Learning Era*, 105 GEO. L.J. 1147, 1160–1, 1167 (2017) (discussing black box problems);

is inversely related to system accuracy (or other performance objectives).[62] Nevertheless, we still must use human experts to explain machine data to the trier of fact to achieve sufficient understanding and trustworthy fact-finding. This is the case despite the fact that the machine is the superior "expert" when it comes to accurate data collection and evaluation.

The rationale of using experts to improve the trustworthiness in the establishment of facts where the trier of fact lacks the relevant knowledge[63] is the same in both the adversarial and inquisitorial systems, but the means of doing so differ widely. In an adversarial system, expert witnesses are typically called by parties, based upon their certification, skills, or experience, to testify before a judge or jury to assist their case.[64] In an inquisitorial system, expertise is sought by the bench where they have determined that they lack the relevant knowledge;[65] expert testimony is generally given orally during the public hearing but can also be provided in written reports.

Regardless of whether the fact-finders are jurors or members of a bench, understanding the issues around the reliability of devices that autonomously make assessments and that may (or may not) be useful in reconstructing the facts of a case exceeds the knowledge and understanding of an average human. As such, the use of complex technology in fact-finding makes expert evidence crucial in both adversarial and inquisitorial justice systems.

b. AI-Generated Machine Evidence

The focus of this Article is machine data generated by a consumer product during an AI-driven interaction. Related to this, Andrea Roth coined the term "machine testimony" or machine evidence to distinguish mere *tools* that assist humans in conveying information from *intelligent machines* that can convey a message of their own by registering

see also FRANK PASQUALE, THE BLACK BOX SOCIETY: THE SECRET ALGORITHMS THAT CONTROL MONEY AND INFORMATION (2015) (providing a more detailed discussion).

[62] *See* Doshi-Velez & Kortz, *supra* note 60, at 2.

[63] *See* FED. R. EVID. 702. *But see* Michael H. Graham, *Expert Witness Testimony and the Federal Rules of Evidence: Insuring Adequate Assurance of Trustworthiness*, 1986 U. ILL. L. REV. 43 (1986); BRUCE D. SALES & DANIEL W. SHUMAN, EXPERTS IN COURT: RECONCILING LAW, SCIENCE, AND PROFESSIONAL KNOWLEDGE (2005) (discussing the ongoing issue of expert trustworthiness). With a specific focus on digital tools*:* Jennifer N. Mellon, *Manufacturing Convictions: Why Defendants Are Entitled to the Data Underlying Forensic DNA Kits*, 51 DUKE L.J. 1097, 1101 (2001).

[64] FED. R. EVID. 706.

[65] The bench is authorized to appoint a "neutral" expert in appropriate cases. Strafprozeßordnung [StPO] [German Code of Criminal Procedure] as amended Apr. 7, 1987, § 73 [hereinafter StPO].

and assessing specific data based upon the device's design and algorithms.[66] The underlying question is whether the same safeguards that are in place for human statements should apply when machine data is offered as truth of the matter asserted. This would include something equivalent to the right to confront (and impeach) a witness and the exclusion of associated evidence where the witness cannot be adequately confronted.

This debate primarily focuses on forensic instruments used in criminal litigation—that is, tools that produce evidence subsequently offered as fact. The issues around the applicability of digital tools engineered for forensic use are different and not the focus of this Article. In the case of machine-generated data, it is produced without regard for criminal proceedings and, more importantly, includes observations and assessments of humans by machines.

In the case of drowsiness detection systems, data from various sources is registered, each of which can act as a separate piece of evidence. Such data can include things like observation of facial features, an assessment that a driver is sleepy, deployment of an alert to the driver, and the driver's response to the alert. While some of this data does not convey an assessment by the machine, a large portion of it does involve evaluation by software bots and the line between traditional tools and robots becomes blurred. With regard to fact-finding in criminal proceedings, addressing the issue of trustworthiness in evidentiary production will be crucial to distinguish tools from the source delivering the message.

c. Consumer Products Generating Machine Evidence

Digital interfaces linking humans and robots are regularly designed as part of a technological solution for a consumer product, such as automated driving, which offers vast possibilities if they could also be used for law enforcement purposes.

However, tapping into the potential of AI-driven devices also raises a multitude of evidentiary issues and a number of problems for establishing facts in criminal proceedings, some of which have already been mentioned (see *supra* II.C.2.a). Robots and software bots offer an almost limitless and indefatigable capacity to register information in their environment and can provide data beyond simple measurements as a result of their ability to continuously record, assess, and document

[66] *See* Roth, *supra* note 3, at 1979, 1986–93, 2002–22; Roth, *supra* note 30, at 1301 (coining the terms "machine testimony" and "automated proof"); *cf.* Chessman, *supra* note 5, at 197, 183, 206, 222 (using the phrase "evidence created by computer programs").

human behavior. In the case of automated driving, this constant monitoring provides a large amount of data to support the determination of a human's fitness to drive.

Despite the capacity to collect vast amounts of data, AI-driven devices cannot explain for themselves how they evaluate human conduct or reach a decision. Therefore, law enforcement and the courts must be cautious about what they learn from machine-generated data.

d. Meeting the Evidentiary Challenge

The previous explanations illustrate that machine evidence is a challenge in many respects. First, the information generated by AI is stored digitally and must be retrieved and subsequently interpreted by an expert. Second, and possibly more important, is the issue that using AI to assess a driver's alertness could be interpreted as a professional statement and may not be explainable in detail because human comprehension is limited by the black box problem.[67] Despite the presence of these issues, AI may take on the role of an eye-witness and implicate a defendant in wrongdoing. Concern that triers of fact will place unyielding trust in such statements, as is occasionally the case with eyewitnesses, seems warranted.

To date, the specific means by which machine evidence can be reliably translated into digestible information are unclear and the relevant admissibility standards remain unresolved.[68] The extent of the black box problem appears to be directly related to the accuracy of the information generated by robots, which limits experts' explanations and other means of testing trustworthiness. The fundamental differences between machine-generated evidence and more traditional types of evidence renders typical means of scrutinizing AI statements impossible. Unlike human witnesses, neither robots nor software programs can be put on the witness stand and asked to take an oath to tell the truth. They are also not deterred from lying by the threat of being prosecuted for perjury. Despite all these problems, machine evidence generated by AI during collaborative actions with humans still holds the promise of vast amounts of information potentially relevant to criminal investigations, especially with human-robot interaction on the rise.

[67] Daniel J. Grimm, *The Dark Data Quandary*, 68 AM. U. L. REV. 761, 819 (2019).

[68] See Jennifer L. Mnookin, *Of Black Boxes, Instruments, and Experts: Testing the Validity of Forensic Science*, 5 EPISTEME 343 (2008); Bathaee, *supra* note 6.

3. The Evidentiary Cycle and Consumer Products Assisting Law Enforcement

Case law[69] and scholarship[70] suggest a predictable life cycle for many types of new evidence, beginning with the assumption that it is initially *too new to be reliable*. It then becomes *new but subject to testing* and then *generally reliable but occasionally improperly applied*. Finally, many types of evidence reach a point of *being blindly trusted*.[71] With the benefit of hindsight, we know this evidentiary life cycle is not irreversible—DNA-testing, for instance, was once blindly trusted but is now under increased scrutiny.[72] That said, reversing the evidentiary cycle is an uphill battle and one that is often preceded by a great deal of human suffering. Therefore, the issue of the initial admissibility of machine evidence becomes of particular importance when it is proffered as a potential type of third-generation forensic evidence. This is because the evidence in question is automatically produced from an AI-driven human-robot interaction through the use of consumer products.[73] While seemingly objective, this evidence might be prone to error, and must still be explained (at least in part) through the use of experts.[74]

Most of us today believe that increased use of machine data in fact-finding, including forensic (e.g., DNA-testing) and non-forensic (e.g., GPS tracking) technology, has resulted in an overall increase in

[69] *See, e.g.*, United States v. Beasley, 102 F.3d 1440, 1448 (8th Cir. 1996); US v. McCluskey, No. 10-2735 JCH, 2013 WL 1239717 at 2 (D. N.M July 2, 2013); Texas v. Josiah Sutton (District Court of Harris County, Cause No. 800450) (2003); People v. Castro, 545 N.Y.S.2d 985, 987 (N.Y. Sup. Ct. 1989); State v. Butterfield, 27 P.3d 1133, 1143 (Utah 2001); Spencer v. Commonwealth, 384 S. E.2d 785, 797–98 (Va. 1989).

[70] *See* JAY D. ARONSON, GENETIC WITNESS: SCIENCE, LAW AND CONTROVERSY IN THE MAKING OF DNA PROFILING (Rutgers University Press 2007).

[71] This observation has been made by Richard Myers. Richard Myers, Remarks at the Data, Technology and Criminal Law Workshop at Duke University (April 5–6, 2019); *see also* United States v. Beasley, 102 F.3d 1440, 1448 (8th Cir. 1996) (holding that "the reliability of the PCR method of DNA analysis is sufficiently well established to permit the courts of this circuit to take judicial notice of it in future cases . . . " but it remains to be seen when this method will be questioned again).

[72] Frederika A. Kaestle, Ricky A. Kittles, Andrea L. Roth & Edward J. Ungvarsky, *Database Limitations on the Evidentiary Value of Forensic Mitochondrial DNA Evidence*, 43 AM. CRIM. L. Q. 53, 85–87 (2006).

[73] Roth, *supra* note 3, at 1975.

[74] *See infra* Part III.B.5. Again, for the purposes of this Article, it is irrelevant whether the data would be proffered as direct evidence (of the fact that the TOR-request has been launched) or as circumstantial evidence (of sleepiness if the drowsiness detection system observed body signals that it registered as signs of driver fatigue).

accuracy and objectivity in reconstructing the facts of a case.[75] However, beyond the immense privacy concerns, this assumption carries with it the risk of blindly trusting machine accuracy and is counterintuitive given that most people, including judges and jurors in criminal cases, do not understand the underlying technology.

While scholars in adversarial systems (especially in the United States) are increasingly denouncing blind faith in this opaque machinery, citing erroneous breathalyzers used in the 1960s and misidentifications through DNA tests in the 2000s, an equally pronounced debate has not surfaced in Europe. Additionally, despite the available scholarship in the United States, the majority of U.S. courts are hesitant to heed requests to scrutinize evidence the public has already deemed safe.[76]

The implication here is that the threshold for relevance and probative value, which, once met, results in a presumption of admissibility for *useful* evidence (as opposed to evidence designed to mislead and confuse the factfinder), may need to be elevated. If trust has been established by the smooth operation of an AI-driven system, say an automated digital driving assistant, judges may be especially skeptical when a defendant attempts to challenge its supposed flawless functioning during a criminal trial. This issue may be related to the fact that judges, like anyone else, are not capable of possessing an expert degree of knowledge on every topic that enters their courtroom. While understandable, this limitation may also result in prohibiting the defense from effectively challenging the relevance of certain expert evidence under the auspices that such a challenge has a high likelihood of misleading jurors.[77]

It is unclear how this general concern will translate in today's digital era. It is likely that the issue of evaluating the trustworthiness of machine evidence will become more pressing with an increase in AI-driven consumer products, like automated cars. Along with an increase in product volume and popularity, trustworthiness will also be important as machine evidence moves through the evidentiary life cycle, particularly from the phase of *too new to be reliable* to *new but subject to testing*.

There are a number of reasons that the current state of machine evidence remains in the *too new* phase. First, there have been no certification processes to date to support the reliability of the data generated

[75] Roth, *supra* note 3, at 1975–76.

[76] *See, e.g.*, Edward J. Imwinkelried, *Computer Source Code: A Source of the Growing Controversy over the Reliability of Automated Forensic Techniques*, 66 DEPAUL L. REV. 97, 99–101 (2016).

[77] *See, e.g.*, Kaestle, *supra* note 72, at 81–86 (discussing DNA evidence); Celentino, *supra* note 45, at 1325–30 (discussing Facial Recognition Technology).

for use in criminal trials because the products have not been designed to accurately record specific data for use in fact-finding in a criminal trial, but rather to meet a broader consumer demand. Further, the data is owned by private individuals or corporations and stored in cloud-based systems where it may also be encrypted and shielded by manufacturers claiming trade secret privileges.[78]

At this point, governments have little to no information about the reliability of any item of machine evidence generated by a consumer product. In the case of drowsiness detection systems, unbeknownst to the driver, the creators of the AI monitoring human behavior may be inclined to design the algorithms in such a way that it shifts blame from the car to the human[79] to protect the corporate self-interest described above.

There are many ways in which a robot's output can be imprecise or ambiguous, including human error at the programming stage, biased algorithms, or biased standardization data, just to name a few. Relative to government devices that are more regulated, consumer products that generate data may be more likely to have hidden (potentially unintentional) subjectivities. Even assuming that a manufacturer using narrow AI would strive for optimal neutrality of machine evidence, computer engineers unintentionally (and unavoidably) create biases.[80] In the case of drowsiness detection systems, the choice of a particular design to capture a driver's face or body position and the trade-offs given to achieve functionality could have dire consequences. For example, variations in eyelid positioning across ethnicities would need to be accounted for so as not to erroneously interpret individual variations as a sign of sleepiness. Such safeguards must be put into place to ensure a certain degree of transparency and authenticity.[81]

[78] *See* Rebecca Wexler, *Life, Liberty, and Trade Secrets: Intellectual Property in the Criminal Justice System*, 70 STAN. L. REV. 1343 (2018) (describing the debate over resolving conflicting interests between a defendant's right to a defense and trade secrets); Natalie Ram, *Innovating Criminal Justice*, 112 NW. U. L. REV. 659 (2017).

[79] *Cf.* Roth, *supra* note 30, at 1272.

[80] *See, e.g.*, CATHY O'NEAL, WEAPONS OF MATH DESTRUCTION (2016) (discussing discriminatory algorithms and their impact on society); *see also* MIREILLE HILDEBRANDT, SMART TECHNOLOGIES AND THE (ENDS) OF LAW: NOVEL ENTANGLEMENTS OF LAW AND TECHNOLOGY 34 (2015); Emily Berman, *A Government of Laws and Not of Machines*, 98 B.U.L. REV. 1277, 1325–27 (2018). For a reflection of the specific application (using the example of facial recognition technology), see GEORGETOWN CTR. ON PRIVACY AND TECH., THE PERPETUAL LINE-UP, https://www.law.georgetown.edu/privacy-technology-center/publications/the-perpetual-line-up.

[81] Berman, *supra* note 80, at 1325–27.

III. A Comparative Perspective of AI in the Courtroom

The following comparative position proposes significant changes to the German and U.S. criminal justice systems with respect to machine-generated evidence. While Germany must strengthen the legal tools available to defendants, the United States needs to continue to broaden the defense's access to forensic evidence by allowing out-of-court statements and reports by experts to be admitted and shared with both parties in order to optimize the objective scientific evidence provided to the trier of fact.

A. *In Pursuit of the Truth*

The primary goal of fact-finding in any criminal proceeding is to establish the truth.[82] In Germany and the United States, there is strong public interest in determining the truth with the hope that it is on the basis of "true" facts that courts make decisions of guilt or innocence.[83] The fact that both systems today conclude most criminal proceedings with some type of plea bargaining is not necessarily opposed to their truth-seeking commitment as both systems operate under the assumption that a defendant's confession reveals the truth.[84]

Certainly, both legal systems share similar rules around the foundation required for trustworthy fact-finding, including independent and impartial judges and formal requirements around evidentiary proceedings. Despite this shared foundation, adversarial and inquisitorial trials

[82] StPO § 244, para. 2 (stating that "[i]n order to establish the truth, the court shall, *proprio motu*, extend the taking of evidence to all facts and means of proof relevant to the decision"); Thomas Weigend, *Should We Search for the Truth, and Who Should Do it?*, 36 N.C. J. INT'L L.& COM. REG. 389, 389 (2011). However, neither the U.S. Constitution nor federal law expressly require investigators, prosecutors, or courts to seek the truth. Seemingly to the contrary, the Federal Rules of Criminal Procedure suggest that certain principles should guide interpretation "to provide for the just determination of every criminal proceeding, to secure simplicity in procedure and fairness in administration, and to eliminate unjustifiable expense and delay." FED. R. CRIM. PROC. R. 2. Case law does, however, point to the search for truth as an underlying guiding principle in criminal proceedings. *See, e.g.*, Tehan v. United States, 383 U.S. 406, 416 (1966); United States v. Havens, 446 U.S. 620, 626 (1980); Colorado v. Connelly, 479 U.S. 157, 166 (1986). For a more detailed discussion, see Jenia Iontcheva Turner, *The Exclusionary Rule as a Symbol of the Rule of Law*, 67 SMU L. REV. 821, 829 (2014).

[83] Jenia Iontcheva Turner & Thomas Weigend, *The Purposes and Functions of Exclusionary Rules: A Comparative Overview*, *in* DO EXCLUSIONARY RULES ENSURE A FAIR TRIAL? 255, 260 (Sabine Gless & Thomas Richter eds., 2019).

[84] Thomas Weigend, *The Potential to Secure a Fair Trial Through Evidence Exclusion: A German Perspective*, *in* DO EXCLUSIONARY RULES ENSURE A FAIR TRIAL? 61, 64 (Sabine Gless & Thomas Richter eds., 2019).

use very different procedural approaches.[85] These differences result most notably from a divergence in the fact-finding body, which is a bench comprised of judges and laypeople in the inquisitorial system and a judge or jury in the adversarial system. This disparity shapes evidentiary rules, including how statements are used in establishing facts and the necessity of a reasoned verdict that can withstand an appeal.[86] Both jurisdictions' common goal of pursuing the truth (albeit procedurally different), combined with their apparent endorsement of automated driving and other AI-driven devices, results in an interesting comparative study.

B. *AI in Adversarial and Inquisitive Courtrooms*

As humans have been increasingly willing to interact with technology and AI-driven devices in recent years, the opportunity to monitor their behavior has vastly increased. The resulting machine evidence may potentially enhance fact-finding, but at the moment criminal justice systems around the world are not yet equipped to adequately handle such data. They lack specific tools to thoroughly vet its reliability or validity and both inquisitorial and adversarial systems must use experts to explain such evidence to the trier of fact as it cannot be grasped with the naked eye.

1. Machine Evidence in Modern Day Courtrooms

Machine evidence does not fit into the conventional evidentiary and procedural scheme whereby humans communicate with each other in a formalized way in pursuit of the truth. Therefore, one must either create an entirely new model of evaluating the reliability of machine evidence in criminal proceedings or rethink the available types of evidence and differentiate between possible systemic and judicial weaknesses to see if judges, prosecutors, and defense attorneys can somehow adequately examine such evidence within the framework of their legal system.[87]

Experts are crucial to the use of machine evidence in a criminal trial. They must capture and clarify how particular data is registered in addition to explaining the impact of a particular machine learning device and its possible sources of error relevant to fact-finding. In an adversarial proceeding, expert evidence is commonly used as part of the

[85] Weigend, *supra* note 59, at 253–265.

[86] Damaška, *supra* note 14, at 426.

[87] Myers, *supra* note 71.

partisan presentation of a case, whereas during an investigation in an inquisitive system, the prosecutor will typically commission experts and subsequently add their reports to the case file. It is important to note that these reports tend to describe methods and explain results, but lack information about how raw data is measured or how a digital evidentiary tool is set up and used.[88] Thus, before the defense can raise an argument about how different forms of evidence should or should not enter fact-finding, a certain narrative of the case has already been put into place via the case file. If an evidentiary report is in the file, the bench may choose to admit the expert evidence based on the written report, by calling the expert to testify, or by calling in another expert to submit a new report.[89] The bench may also anticipate the need for further fact gathering and summon other experts at the oral trial hearing.[90]

In Germany, defense attorneys have access to the entire case file by the end of the pretrial investigation and have the ability to ask the court to summon an expert to appear at the trial so that he or she may be questioned. Where they have serious doubt about an expert's credibility, they may bring their own expert to trial (provided resources are available), but the bench, as the driving force behind fact-finding, is not always required to hear such testimony. German law warrants a rejection if, from the point of view of the bench, the alleged fact in question has already been clearly proven (or disproven) by the first expert opinion.[91] Notably, this rule does not apply to cases "where the professional competence of the first expert is in doubt, where her opinion is based upon incorrect factual suppositions, where the opinion contains contradictions, or where the new expert has means of research at his disposal which seem to be superior to the ones of an earlier expert."[92] This statutory provision is applied differently depending on where in the evidentiary cycle a piece of evidence is; a novel forensic technique is more likely to result in the bench allowing the opinion of more than one expert, whereas a generally reliable, or even blindly

[88] *See, e.g.*, Oberlandesgericht [OLG] Bamberg, June 13, 2018, 3 Ss Owi 626/18 (Ger.); Oberlandesgericht [OLG] Karlsruhe, July 16, 2019, 1 Rb10 Ss 291/19 (Ger.), Verfassungsgerichtshof [VERFGH] Saarbrücken, Apr. 27, 2018, Lv 1/18 (Ger.); Kammergericht [KG] Apr. 2, 2019, 3 Ws [B] 97/19 – 122 Ss 43/19 (Ger.) (addressing radar guns used to detect speeding drivers).

[89] StPO § 221, § 222, § 244, § 256.

[90] *Id.* § 245.

[91] *Id.* § 244, para. 4.

[92] *Id.*

trusted forensic tool will face greater challenges where a second expert's opinion is requested.

German procedural law developed in the nineteenth century, but has been, at least to some extent, influenced by adversarial notions since the 1950s as a result of case law from the European Court of Human Rights (ECtHR), the prominent human rights tribunal based on the European Convention on Human Rights (ECHR).[93] In particular, the notion of a fair trial, including the right to examine incriminating evidence (Art. 6 ECHR) has had a lasting effect on fact-finding in Continental Europe and often serves as a sort of backup if the traditional inquisitorial system lacks adequate protection for the individual.[94] While the idea of challenging the reliability of machine evidence using the right to examine an incriminating witness under Article 6, paragraph 3(d) of the ECHR[95] is unlikely to be embraced by German courts anytime soon,[96] recent case law from Higher Regional Courts (i.e., the highest courts in any given state) does suggest that an increasing number of judges may be open to the idea of allowing access to so-called "raw measure data" in order to more thoroughly vet machine evidence like digital radar guns.[97] These decisions seek to achieve "knowledge parity" in an effort to meet the benchmark of European case law on Art. 6 ECHR and its "equality of arms" between the prosecution and defense and to strengthen the defense's position with regard to the bench.[98] This concept, while seemingly adversarial, has been said to be grounded in both civil and common law traditions and is a consequence of the ECtHR's attempt to create a cross-jurisdictional notion of procedural fairness.[99]

This new line of argument, in some ways, parallels the call by scholars in the United States that machine evidence be viewed as out-of-court

[93] *See* Roberto E. Kostoris, *European Law and Criminal Justice*, *in* HANDBOOK ON EUROPEAN CRIMINAL PROCEDURE 47–56 (Roberto E. Kostoris ed., 2018) (detailing ECtHR's jurisprudence and binding effect on Member State courts); Weigend, *supra* note 82, at 64 (describing the effect in Germany specifically).

[94] *See generally* JOHN D. JACKSON & SARAH J. SUMMERS, THE INTERNATIONALISATION OF CRIMINAL EVIDENCE: BEYOND THE COMMON LAW AND CIVIL LAW TRADITIONS 79–95 (2012).

[95] Convention for the Protection of Human Rights and Fundamental Freedoms art. 6, Nov. 4, 1950, 213 U.N.T.S. 222.

[96] *Cf.* Oberlandesgericht [OLG] Bamberg, June 13, 2018, 3 Ss Owi 626/18 (Ger).

[97] *Cf.* Oberlandesgericht [OLG] Karlsruhe, July 16, 2019, 1 Rb10 Ss 291/19 (Ger.); Verfassungsgerichtshof [VERFGH] Saarbrücken, Apr. 27, 2018, Lv 1/18 (granting access to measurement data based on Article 6 of the European Human Rights Convention).

[98] Jürgen Cierniak & Holger Niehaus, *Neuere Entwicklungen im Recht auf Einsichtnahme in Messunterlagen*, 14 DEUTSCHES AUTORECHT [DAR] 541, 541–44 (2018) (Ger.).

[99] *See* Jackson & Summers, *supra* note 94, at 79–80.

testimony offered as truth of the matter asserted (and in need of context).[100] Even if one agrees with those who think that AI-driven devices should undergo similar credibility testing as witnesses because of their design, standardization data, or machine learning software, one must also be aware that should this argument be accepted, it could improperly place such machines on similar footing as human witnesses.

To illustrate this point, if a human passenger in a car was put on the stand to testify about a defendant's driving ability, he or she would be questioned about perceptual capacities, potential biases, misjudgment, or even intentional lying (with the risk of prosecution and punishment for perjury). As of today, AI-driven devices cannot undergo the equivalent of cross-examination even where they are evaluating human users and coming to a conclusion, like whether or not a driver has the capacity to operate a vehicle. If such determinations by AI are used as evidence, it should be subject to scrutiny, especially with respect to the design, algorithms, and machine learning/training data.

This becomes exceptionally complicated given that a thorough evaluation and understanding of the inner workings of an AI-driven device can only happen outside the courtroom because of the degree of complexity and the desire for the corporate world to protect their trade secrets. As a result, in an adversarial system, this evaluation would not be able to be included in a hearing. Instead, experts would appear in court to speak about the results of the data retrieved and act as a sort of proxy for direct contact with the particular device regarding its reliability and validity.

2. Testing Machine Evidence for Relevance and Reliability

Only relevant and reliable evidence can be presented in courts in Germany or the United States. Although German law lacks an explicit blanket rule laying out the requirements of admissibility or how to determine the reliability of evidence, courts follow the principle that all relevant evidence is to be admitted as a natural part of their truth-seeking mission.[101]

Despite the fact that this principle is also enshrined in common law, the U.S. Federal Rules of Evidence memorialize it in Rules 401 and 402.

[100] *See* Joëlle Vuille, Luca Lupària & Franco Taroni, *Scientific Evidence and the Right to a Fair Trial under Article 6 ECHR*, 16 L. PROBABILITY AND RISK 55 (2017); Paul Roberts & Michael Stockdale, *Introduction: Forensic Science, Evidential Reliability and Institutional Reform, in* FORENSIC SCIENCE EVIDENCE AND EXPERT WITNESS TESTIMONY: RELIABILITY THROUGH REFORM? 22 (Paul Roberts & Michael Stockdale eds., 2018).

[101] *See* StPO § 261; *see also* Weigend, *supra* note 82, at 389.

While only relevant evidence may be admitted, not all relevant evidence is admissible. Evidence is relevant if there is a particular connection between it and the fact it is offered to prove or disprove. This connection does not have to be so strong that a single item of evidence has any tendency to make a fact more or less probable.[102] It is sufficient if the piece of evidence amounts to a link in a chain of information offered as proof. Data gathered as a result of monitoring a human driver's face or the fact that a drowsiness alert was activated could be relevant evidence, but so could an overall assessment of a driver's conduct by a driving assistant should it make a material fact more or less probable than if it were excluded.

There are many rules that can lead to the exclusion of potentially relevant evidence, including Rule 403, which authorizes judges to balance the probative value of an item of evidence against the potential harm resulting from its admission. The same rule states that relevant evidence should be excluded where its probative value is outweighed by the danger of unfair prejudice, confusion to the trier of fact, or where it is deemed to be wasting time or cumulative in nature. Doubt surrounding the source of a piece of evidence can also lead to the evidence being viewed as less credible.[103] Therefore, if a judge determines that the reason for which a drowsiness detection system in a car was triggered is not of sufficient probative value to determine whether or not the driver was sleepy, he or she will exclude the evidence. Although courts tend to interpret Rule 403 narrowly,[104] a great deal depends on the judge's reasoning, particularly with respect to whether or not the evidence would confuse or mislead a jury.[105] This issue is especially pertinent to the black box problem inherent in machine-generated data given the high degree of technicality and limited means of explaining it.[106]

In addition to the requirement that evidence be relevant, it must also be reliable. For example, polygraph evidence lacks reliability and, as a result, is generally banned from federal courtrooms in America.[107] In

[102] *See* SANFORD H. KADISH, STEPHEN J. SCHULHOFER & RACHEL E. BARKOW, CRIMINAL LAW AND ITS PROCESSES: CASES AND MATERIALS 25–26 (10th ed. 2017).

[103] *See* Edward J. Imwinkelried, *The Meaning of Probative Value and Prejudice in Federal Rule of Evidence 403: Can Rule 403 Be Used to Resurrect the Common Law of Evidence?* 41 VAND. L. REV. 879, 881–86 (1988) (describing additional information about Rule 403).

[104] *Id.* at 884, 886–89.

[105] John Nawara, *Machine Learning*: *Face Recognition Technology Evidence in Criminal Trials*, 49 U. LOUISVILLE L. REV. 601, 607 (2010).

[106] *See supra* Part II.C.2.a.

[107] The United States Court of Appeals for the Fourth Circuit generally bars them and the U.S. Attorneys' Manual for prosecutors suggests that Assistant U.S. Attorneys should oppose their

the words of Justice Thomas in *Scheffer*, "there is simply no consensus that polygraph evidence is reliable[.]"[108] However, some federal appellate courts have abandoned this per se exclusionary rule and have left the decision of admission or exclusion to the discretion of district courts under *Daubert*,[109] thereby granting more leeway with regard to the relevance and reliability of new technologies in the future.[110] It may be the case that judges' attitudes will continue to change as the AI technology becomes increasingly useful and more ubiquitous.

Given the reticence of European courts to use lie detectors, one might doubt a similar scenario would come to fruition anytime soon. In 1998, the highest German Court (Bundesgerichtshof) deemed polygraph evidence as "a completely unsuitable means of proof" without any probative value.[111] It further held that the polygraph's measurement of bodily functions, specifically the registration and assessment of data, lacks sufficient scientific methodology to be considered reliable evidence.[112]

Despite the treatment of polygraph evidence by courts around the world, technology continues to develop and the use of processes such as AI-driven Facial Recognition Technology and other tools that use machines to monitor and evaluate human behavior have increased. As such, criminal courts may begin to have more difficulty arguing that machine-generated evidence is not adequately equipped to assess the mental state of a person or predict human action. With respect to drowsiness detection systems, evidence has shown that they predict

introduction as unreliable. *See* U.S. DEP'T OF JUSTICE, U.S. ATTORNEYS' MANUAL: CRIMINAL RESOURCE MANUAL § 262, https://www.justice.gov/usam/criminal-resource-manual-262-polygraphs-introduction-trial.

[108] United States v. Scheffer, 523 U.S. 303 (1998). To this day, the scientific community worldwide remains extremely polarized regarding the reliability of polygraph techniques. *See, e.g.*, Jacqueline Elton, *The Polygraph in the English Courts: A Creeping Inevitability or a Step too Far?* 81 J. CRIM. L. 66, 68–74 (2017); *see also* United States v. Posado, 57 F.3d 428, 434 (5th Cir. 1995); United States v. Cordoba, 104 F.3d 225, 228 (9th Cir. 1997) (for U.S. examples). At least one federal appellate court has recently reaffirmed its *per se* ban. *See* United States v. Sanchez, 118 F.3d 192, 197 (4th Cir. 1997). The Second Circuit recently noted that it has "not decided whether polygraphy has reached a sufficient state of reliability to be admissible." United States v. Messina, 131 F.3d 36, 42 (2d Cir. 1997).

[109] Daubert v. Merrell Dow Pharm., Inc., 509 U.S. 579 (1993); David E. Bernstein & Jeffrey D. Jackson, *The Daubert Trilogy in the States*, 44 JURIMETRICS 351 (2004) (explaining *Daubert* in detail).

[110] Nawara, *supra* note 105, at 605.

[111] Bundesgerichtshof [BGH] [Federal Court of Justice] Dec. 17, 1998, 1 StR 1998, 156/98 (Ger.).

[112] *Id.* at 44–74.

momentary episodes of somnolence quite well,[113] and it is for this reason that such technology will become part of EU vehicle safety measures. It may therefore be the case that machine evidence will enter the evidentiary life cycle in Europe in the not-too-distant future.[114]

3. Use of Written Reports to Introduce Machine Evidence

If machine evidence was determined to be sufficiently reliable, could it then be presented at trial in the form of written reports submitted by experts (as opposed to introduction by oral testimony)? Or would such reports be excluded because a finding by a machine was offered as truth of the matter asserted, thereby triggering confrontation rights, including the testimonial safeguards of impeachment and hearsay rules? In Germany such reports would, in principle, be acceptable as long as both parties have trust in the court-appointed expert and have full access to the report via the shared case file. In the United States, however, the presentation of such reports would likely meet considerable resistance in light of a long tradition of an in-court evaluation of evidentiary credibility and a broad interpretation of what constitutes "testimony."

a. Germany

In Germany, it is normal operating procedure in a criminal trial for the prosecution to include a lab report in the case file early in the proceedings. Such a document would state that an expert was appointed by the prosecution service or the bench to evaluate the merits of the case and would include the tests administered and subsequent findings by the expert. The lab report generally will not, however, reveal details about a digital measuring device like a radar gun. As long as the device has been certified to be used as an evidentiary tool, the measured raw data and the details of the digital design will typically not be disclosed to the defendant. For example, a report in a case where a drowsiness detection system's data is offered as evidence could include things like if and when a driver was alerted that they were too tired to drive and any changes in the intensity of the alert over time as calculated by the expert. As things stand today, the report would not include details like the software design, methods used in machine learning, or the machine training data.

[113] Vural et al., *supra* note 9 (claiming that their system predicted crashes related to sleepiness during a driving computer game with 96% accuracy).

[114] Kaestle, *supra* note 72, at 53 (describing such a cycle with regard to DNA evidence).

The trier of fact (the bench in this case) may accept these types of reports as expert evidence asserting as fact that, for instance, the drowsiness detections system's coffee cup-signal was illuminated, but also as corroborative evidence that the driver was tired and did not stop for a break, which could amount to a failure to exercise due care in the circumstances or may even reach the point of recklessness. The bench, as the natural driver of fact-finding in an inquisitorial proceeding, may also choose to appoint another expert to provide a new report or summon further experts to assess the machine's findings in the oral hearing.[115] Presently, it is unclear how much tinkering or evaluation of an AI-driven device would be allowed or deemed necessary to establish "the truth." It is, however, clear that a court-appointed expert has sufficient time and, at least theoretically, powerful means to seize data through the help of the prosecutor or the bench that appointed her or him.

During the investigation phase, a defendant can informally offer expert evidence to support his or her claim(s) by giving the lead prosecutor a report from an expert. This is, of course, dependent upon the defendant having the financial resources to hire an expert as well as access to the necessary case information.[116] If such a report was generated, it would be added to the case file by the prosecution and the defense would then be contributing their own witness's opinion. All documents in the case file then become part of the court's truth-seeking mission and the file is accessible to all parties before the case goes to trial.[117]

The inquisitorial justice process is built upon the general assumption that the bench, as professionals, are impartial at the outset of every case and are sufficiently experienced to identify unreliable evidence. With regard to forensic evidence, a great deal of trust is placed in governmental institutions that work closely with the prosecution, who is obliged to look for both incriminating and exonerating evidence.[118] Not without good reason, Mirjan Damaška has argued that fact-finding by the bench and the traditionally episodic and placid approach on the European continent is likely the primary reason for an absence of clear evidentiary rules, and it seems more than likely that this problem will

[115] StPO § 214.

[116] *Id.* § 222.

[117] *Id.* § 261.

[118] *Id.* § 256, para. 1; Matthias Krüger, *Commentary StPO § 256*, *in* MÜNCHENER KOMMENTAR ZUR STRAFPROZEßORDNUNG BD.2: §§ 151–332 (Kudlich et al. eds., 2016) (Ger.).

persist in the future.[119] However, the generous amount of time allowed to vet experts outside the courtroom is a particularly important aspect of the inquisitorial system and one could also argue that the lack of clear evidentiary rules provides leeway for a more flexible approach.

Nevertheless, such notions of the open-minded bench and impartial law-enforcement clearly are idealistic and place a significant amount of faith in the state. A supposed safety net is traditionally created through a review of the establishment of facts by appellate courts.[120] As such, in practice, benches administer proceedings with the possibility of appellate review in the back of their minds, which makes the evidentiary process predictable despite a lack of strict evidentiary rules. The corresponding obligation for the bench to explain the reasoning behind the evaluation of evidence in a judgment can lead to individual and comprehensive explanations that render the process transparent, but may also result in cookie-cutter decisions that will hold up on appeal but do not provide much consideration for the specifics of a particular case.[121]

The path machine evidence will take in this system is somewhat difficult to foresee given the array of options for obtaining evidence during the investigatory phases. Much will depend on the prosecution services who set the blueprint for fact-finding very early on via the case file, as well as on the benches that eventually decide how much vetting a piece of evidence requires before it can be deemed reliable in the establishment of the truth. It is interesting to note that over the last few years certified digital evidentiary tools have sparked a heated debate around the scope of access to the file or, more generally, the idea of "knowledge parity."[122] As things stand today, the prosecution or the bench simply adds the expert's report to the file, which is accessible by the defense, but does not include detailed information about how the digital tool works, nor does it provide the defense with any means of collecting or accessing such information. This shortfall has recently been

[119] Damaška, *supra* note 14, at 428–29.

[120] Andreas Mosbacher, *Das Ideal richterlicher Wahrheitsfindung und die Betrübnisse des wirklichen Lebens: Richterliche Schuldfeststellung und die Gefahr des Fehlurteils*, 9 FORENSISCHE PSYCHIATRIE, PSYCHOLOGIE, KRIMINOLOGIE 82, 86 (2015) (Ger.).

[121] Wolfgang Frisch, *Beweiswürdigung und richterliche Überzeugung*, 10 ZEITSCHRIFT FÜR INTERNATIONALE STRAFRECHTSDOGMATIK 707, 711–13 (2016) (Ger.), http://www.zis-online.com/dat/artikel/2016_10_1056.pdf

[122] Rudolf Wendt, Das Recht auf Offenlegung der Messunterlagen im Bussgeldverfahren 30 NEUE ZEITSCHRIFT FÜR VERKEHRSRECHT, 441, 442–43 (2018) (Ger.)

scrutinized in the case law of the Higher Regional Courts,[123] but a decision by the Federal Court on the topic has yet to be made. Nevertheless, it seems likely that in the future machine evidence will be introduced by experts, whether in written reports or orally in more detail than today.

b. United States

In the United States, a report documenting the findings from an expert who is not present in court would meet a great deal more resistance from the defense than it would in Continental Europe.[124] That said, under the current evidentiary regime, courts often rely on legal memoranda and scientific documents rather than oral hearings when examining witnesses in the courtroom.[125] Should this practice also govern machine evidence presented in court? Scholars opposed to this development have gone into great detail explaining how flaws in design (black box problems) and other human/machine errors can lead to unreliable fact-finding when using machine evidence. They liken this to the hearsay dangers hidden in human assertions and thus oppose a documentary evidence approach.[126]

If and when the hearsay rule should apply to reports[127] or whether they (and other documentary evidence) should be admitted as an assertion of fact is the subject of a controversial debate and case law has yet to provide any clarity.[128] This topic is addressed below as part of the

[123] Oberlandesgericht [OLG] Bamberg, June 13, 2018, 3 Ss Owi 626/18 (Ger.). *But see* Oberlandesgericht [OLG] Karlsruhe July 16, 2019, 1 Rb10 Ss 291/19 (Ger.), Verfassungsgerichtshof [VERFGH] Saarbrücken Apr. 27, 2018, Lv 1/18; Kammergericht [KG] April 2, 2019, 3 Ws [B] 97/19 – 122 Ss 43/19 (Ger.) (granting access to measurement data based on Art. 6 of the European Human Rights Convention).

[124] *See, e.g.*, Melendez-Diaz v. Massachusetts, 557 U.S. 305 (2009); Bullcoming v. New Mexico, 564 U.S. 647 (2011); Williams v. Illinois, 567 U.S. 50 (2012); *see also* Celentino, *supra* note 45.

[125] *See, e.g.*, Hayes v. State, 660 So. 2d 257, 262–64 (Fla. 1995); United States v. Porter, 618 A.2d 629, 635 (D.C. 1992); United States v. Havvard, 117 F. Supp. 2d 848, 854 (S.D. Ind. 2000); People v. Palmer, 145 Cal. Rptr. 466, 472 (Cal. Ct. App. 1978).

[126] Roth, *supra* note 3, at 1989–99; Chessman, *supra* note 5, at 209.

[127] *See generally* Madeline Smedley, Note, *Hearsay in the Modern Age: Balancing Practicality and Reliability by Amending Federal Rule of Evidence 801(d)(1)(A)*, 87 GEO. WASH. L. REV. 207 (2019), for additional details about Federal Rule of Evidence 802.

[128] *See* Christopher B. Mueller, Laird C. Kirkpatrick, & Liesa L. Richter, EVIDENCE UNDER THE RULES: TEXT, CASES, AND PROBLEMS 313–19 (9th ed. 2019); Paul C. Giannelli, *The Admissibility of Laboratory Reports in Criminal Trials: The Reliability of Scientific Proof*, 49 OHIO ST. L.J. 671 (1988); *see also* Lyle Denniston, The Confrontation Clause — Again, and Again, SCOTUS BLOG (May. 9, 2014, 2:24 PM), https://www.scotusblog.com/2014/05/the-confrontation-clause-again-and-again/ (last visited January 26, 2020).

discussion of witness evidence (see C.II.3.). Notably, scholars have put forward the argument that where machine evidence is concerned, it may prove more useful to replace a strict requirement for oral hearings with alternative solutions that take into account the complexity of thoroughly testing intelligent machines.[129] This idea may gain traction given that courts and legislatures today tend to be more open to experimenting with evaluations and testimony outside the courtroom.[130]

From a comparative perspective, where machine evidence could be introduced as documentary evidence (e.g., a lab report under an exception to the hearsay rule), its trustworthiness could be assessed through disclosure. In other words, the defense would have the right to know what law enforcement knows about the reliability of the evidence. The out-of-court machine/human assertion admitted under a hearsay exception, could be tested by the defense by arguing that the machine was a declarant. The defense could then use any of the "declarant's" prior (inconsistent) statements to impeach its credibility,[131] including citing the validity testing (or lack thereof) of a drowsiness detection system.

In the case of machine evidence generated by consumer products, it is questionable whether disclosure tools are sufficient to scrutinize its reliability. First, the defense's access to evidence during disclosure is not a particularly powerful right[132] compared to discovery rules in civil cases.[133] Furthermore, remedies provided by case law designed to strengthen disclosure rules fall short of what is needed in the specific situation of machine evidence that is a byproduct of a consumer product, not to mention that the relevant data is stored with private

[129] *See* Maayan Perel & Niva Elkin-Koren, *Black Box Tinkering: Beyond Disclosure in Algorithmic Enforcement*, 69 FLA. L. REV. 181, 185 (2017).

[130] Roth, *supra* note 3, at 2028 (using the phrase "meaningful access to machine evidence").

[131] Roth, *supra* note 3, at 2033.

[132] *See* Ion Meyn, *Discovery and Darkness: The Information Deficit in Criminal Disputes*, 79 BROOK. L. REV. 1091, 1094, 1103 (2014).

[133] Traditionally this imbalance has been justified by privileges and constitutional protections granted to a criminal defendant that theoretically work to his or her advantage, including the right to remain silent or the right to not be required to disclose a defense strategy. *See* United States v. Garsson, 291 F. 646, 649 (S.D.N.Y. 1923) ("Under our criminal procedure the accused has every advantage. While the prosecution is held rigidly to the charge, he need not disclose the barest outline of his defense. He is immune from question or comment on his silence; he cannot be convicted when there is the least fair doubt in the minds of any one of the twelve. Why in addition he should in advance have the whole evidence against him to pick over at his leisure, and make his defense, fairly or foully, I have never been able to see."). However, given the practicality of how these rights work (or don't), they hardly outweigh the prosecutorial advantage in evidentiary discovery.

stakeholders.[134] Primarily based on the Due Process Clause of the Fourteenth Amendment, *Brady v. Maryland*[135] obligates the prosecution to provide the defense with *any* material evidence[136] that would be *reasonably likely* to change the outcome of the trial.[137] Accordingly, in recent years, the forensic evidence held by the prosecution made available to the defense has been considerably expanded.[138]

Despite this obligation, in cases of incriminating machine evidence generated by consumer products, the prosecution will often not be in possession of material that would be likely to undermine the court's confidence in data generated by an automobile's driving assistant. The source code, machine training data, and algorithms used for drowsiness detection systems will typically be in the possession of the car manufacturer, who may refuse to produce them by claiming trade secret privileges.[139] Therefore, defendants seeking usable material to challenge such data cannot rely solely on what is already in the possession of the prosecution. This is different from the situation where a defendant is seeking disclosure of information from digital tools designed for forensic use (like a breathalyzer or DNA sampling) as the Federal Rules of Criminal Procedure mandate the disclosure of such information.[140] Despite this rule, courts have been reluctant to grant discovery of any "underlying documentation" used in preparation of a final report, and specifically any results from digital tools used in forensic settings.[141] Similar issues are likely to arise with AI's use of big data to make

[134] *See* Kerr, *supra* note 47, at 309–10 (providing information on the general issue of cybercrime).

[135] Brady v. Maryland, 373 U.S. 83 (1963).

[136] *Id.* Evidence is deemed "material" if the prosecutor's failure to produce it would undermine the court's confidence in the outcome of the proceedings. Kenneth M. Miller, *Nixon May Have Been Wrong, but it Is Definitely Misunderstood (or, a Federal Criminal Defendant's Pretrial Subpoenas Duces Tecum Properly Reaches Potentially Admissible Evidence)*, 51 WILLAMETTE L. REV. 319, 324 (2015).

[137] *See* Miller, *supra* note 136, 323—26 (for additional information).

[138] U.S. DEP'T OF JUSTICE, JUSTICE MANUAL 9.5.001(F) (2018), https://www.justice.gov/jm/title-9-criminal.

[139] *See* Wexler, *supra* note 78; Ram, *supra* note 78, at 701–4.

[140] *See* Mellon, *supra* note 63, at 1113–14 (arguing that defendants have a right to the source code of digital tools in forensic settings when the term "scientific report" is interpreted to include any information relied upon, either explicitly or implicitly, in creating a final expert report, although most jurisdictions have not embraced this interpretation).

[141] *See* United States v. Iglesias, 881 F.2d 1519, 1524 (9th Cir. 1989) (denying discovery of a chemist's log notes); Roberts v. State, 396 S.E.2d 81, 81 (Ga. Ct. App. 1990) (denying discovery of an expert's notes, work product, recordation of data, internal documents, or graphs); State v. Parnell 883 N.W.2d 652, 667 (Neb. 2016) (disclosure of a cellular analyst's opinion by the State one week before trial did not violate due process).

predictions. It also remains to be seen how courts will react when faced with data generated by consumer products offered as evidence.[142]

Should a defendant seek to challenge the accuracy of a statement from an "intelligent machine," like a drowsiness detection system, he or she might need to access the source code to understand how it was programmed. He or she may also want access to the machine learning algorithms to examine other aspects such as any trade-offs that were made to further efficiency and effectiveness in the AI-driven process.[143] It is against this backdrop that scholars are demanding a "digital *Brady*" rule, or discovery regarding the procedures used to produce a particular set of data, as well as information about their reliability.[144] While the seriousness of this call for action is uncontested, the reality is that most criminal proceedings end in a plea (typically with defendants not entirely understanding the probative value of the machine evidence against them),[145] and therefore the right to impeach a witness.[146] The question of resources is also frequently a determinative factor. Therefore, even with a digital *Brady* disclosure rule in place, it would only be helpful where a defendant could assert his or her right to impeach a witness, even after a plea,[147] and had the access and resources to retain their own expert who could critically evaluate the evidence and ask appropriate questions.[148] This raises numerous concerns about the practicality of the means currently available for indigent defendants to challenge evidence.[149]

[142] A new stand on trade secret privileges would need to be made. *See, e.g.*, Wexler, *supra* note 78.

[143] *See* Berman, *supra* note 80, at 1325.

[144] Garrett, *supra* note 5, at 211–12; *see* Danielle Keats Citron & Frank Pasquale, *The Scored Society: Due Process for Automated Predictions*, 89 WASH. L. REV. 1 (2014); Tom Baker & Benedict G.C. Dellaert, *Regulating Robo Advice Across the Financial Services Industry*, 103 IOWA L. REV. 713 (2018).

[145] Roth, *supra* note 3, at 2033.

[146] *See* United States v. Ruiz, 536 U.S. 622, 629 (2002) (discussing the government's obligations to disclose material impeachment evidence prior to entering a plea agreement); *see also* Colin Miller, *The Right to Evidence of Innocence Before Pleading Guilty*, 53 U.C. DAVIS L. REV. 271, 293–299 (2019) (arguing that under the Due Process Clause a defendant is entitled to evidence of his or her innocence prior to plea bargaining).

[147] *See generally* Miller, *supra* note 146, at 293–99, 320–21.

[148] Murphy, *supra* note 5, at 749.

[149] *See* Stephen A. Saltzburg, *The Duty to Investigate and the Availability of Expert Witnesses*, 86 FORDHAM L. REV. 1709, 1715–20 (2018).

4. Machine Evidence and the Need for Contextualization and Confrontation

Legal safeguards in place in both Germany (e.g., access to the case file, the principle of immediacy, witness attendance obligations, and rights of confrontation) and the United States (e.g., the right to confront and impeach a witness, the rule against hearsay) fall short of assuring valid fact-finding where machine evidence is offered as truth of the matter asserted because both systems lack adequate means to "confront" such evidence in court, thoroughly vet it, and provide context to the trier of fact.

If one assumes that machine evidence introduced by a written report would qualify as pre-recorded testimony offered as truth of the matter asserted, the questions arising thereafter are fundamentally different. Following the Supreme Court's decision in *Crawford v. Washington*, a statement that is considered testimonial in nature may not be introduced at trial against an accused unless he or she has had an opportunity to cross-examine the person who made the statement and that person is unavailable to testify at trial.[150] However, if a statement is not regarded as testimonial, the Confrontation Clause poses little, if any obstacle, to its admission.

The meaning of the word "testimonial," or rather, the type of witness that will trigger the Confrontation Clause, has been the subject of vigorous debates.[151] In Europe, there has been some litigation before the European Court of Human Rights (ECHR) with regard to the right to confrontation under Article 6(3)(d) of the European Convention on Human rights (ECHR).[152] The prevailing doctrine in the United States requires the exclusion of statements by witnesses for the prosecution where the witness is not available to testify at trial unless the defense had a prior chance to challenge such statements by means of cross-examination.[153] The European doctrine is much more vague as the ECtHR currently draws upon a three-step analysis first set out in *Al-Khawaja & Tahery v. U.K.*[154]: (1) whether a good reason exists for the absence of a witness at trial; (2) whether the conviction is solely or

[150] Crawford v. Washington, 541 U.S. 36, 36 (2004).

[151] *See, e.g.*, Richard D. Friedman, *Grappling with the Meaning of Testimonial*, 71 BROOK. L. REV. 241, 242 (2005).

[152] See Vuille, *supra* note 100, for references to applicable case law.

[153] See Denniston, *supra* note 128, for a detailed discussion.

[154] Al-Khawaja & Tahery v. U.K., Eur. Ct. H.R., App. No. 26766/05 & 22228/06, 37 (2011), https://hudoc.echr.coe.int/eng#{%22itemid%22:[%22001-108072%22]}. *Cf.* Jackson & Summers, *supra* note 94, at 93–95, 338–41.

decisively based on the statement of the absent witness; and (3) whether sufficient counterbalancing factors exist "to compensate for the handicaps caused to the defen[s]e as a result of the admission of the untested evidence and to ensure that the trial, judged as a whole, was fair."[155] The ECtHR applied this test in a more recent case involving Germany, emphasizing the trial court's obligation to "approach untested evidence of an absent witness with caution" and look for corroborating evidence,[156] illustrating the unpredictability of this line of case law.[157]

While strict interpretation of the U.S. right may too be waning in practice,[158] scholars maintain that constitutional protections require that a witness take an oath and be confronted by the accused in person while maintaining their assertions under cross-examination.[159] The underlying rationale of preventing false accusations and optimizing the trier of fact's chance to accurately evaluate the credibility of a statement also typically necessitates that the declarant be present in court so his or her demeanor can be taken into account.

This concept appears meaningless in the context of machine evidence because it is impossible for it to undergo a similar vetting process in court. It also does not seem feasible to replace AI-driven devices with one or more of the individuals who created it. This is not just because the formation and operation of AI typically involve a number of individuals, none of whom would be able to fully explain a robot's action where specific machine learning programming was used.[160] When vetting human testimony, we want to know whether certain factors were perceived and considered and how they led the human to a particular conclusion. Artificial Intelligence cannot answer these questions with the technology available today. Although much work has gone into "explainable AI," or the development of machine learning models that are interpretable by humans, as well as the creation of self-explaining

[155] *See* Al-Khawaja & Tahery v. U.K., *supra* note 154, at 37.

[156] Schatschaschwili v. Ger., Eur. Ct. H.R., App. No. 9154/10, 30–31 (2015), https://hudoc.echr.coe.int/eng.

[157] *See* Deborah Paruch, *Testimonial Statements, Reliability, and the Sole or Decisive Evidence Rule: A Comparative Look at the Right of Confrontation in the United States, Canada, and Europe*, 67 CATH. U. L. REV. 136–137, 147–48 (2018).

[158] Strict interpretation of the right was not without initial controversy, as described by Justice Rehnquist, "while I agree that the Framers were mainly concerned about sworn affidavits and depositions, it does not follow that they were similarly concerned about the Court's broader category of testimonial statements . . . " Crawford v. Washington, 541 U.S. 36, 71 (2004) (Rehnquist, J., concurring).

[159] Celentino, *supra* note 45, at 1331.

[160] *See* Chessman, *supra* note 5, at 179.

AI, meaningful advances in explainable decision-making have been minimal.[161] Adequate vetting of AI seems difficult to achieve but will remain a goal given the fact that certain information can be obtained solely from its system. Researchers have continued to work on teaching AI how to make itself more understandable and recently scientists from both private and academic sectors have had some success in teaching image recognition software to show evidence it relied upon in reaching its decision.[162]

Presently, we need the functional equivalent of a means to evaluate AI statements, which would not only involve explanations from human witnesses, but also provide an account of the input, how the data was processed, and the final assessment directly from the machine rather than expert interpretation (just as we would not allow an expert to mediate a human witness's testimony).

a. Germany

The question of whether or not conclusions by AI-driven devices can be evaluated for accuracy similar to human witnesses initially appears to be a moot point under German law as it does not provide for a genuine confrontation of human witnesses and Germany does not have an adversarial-like hearsay rule[163] or a substantial cross-examination process.[164] This is primarily because it is not a jury comprised of laypeople (who presumably require more context) that is assessing the credibility of a source, but rather benches made up of people with trial experience and knowledge of the case file. The German criminal justice system does, however, acknowledge its specific risks, among them a trial based upon a previously prepared case file. It aims to prevent the court's truth-finding mission from being predetermined by the prosecutor's investigation with a commitment to the so-called principle of immediacy: this principle states that the court shall base its judgement only on what has been said and done at the public trial.[165] The inquisitorial

[161] *See* Johannes Fähndrich, Sebastian Ahrndt, & Sahin Albayrak, *Towards Self-Explaining Agents*, *in* TRENDS IN PRACTICAL APPLICATIONS OF AGENTS AND MULTIAGENT SYSTEMS: 11TH INT'L CONF. ON PRACTICAL APPLICATIONS OF AGENTS AND MULTI-AGENT SYS. 147–150 (Javier Bajo Pérez et al. eds., 2013) (for the general idea of a self-explaining system).

[162] *Cf.* Dong Huk Park, et al., *Multimodal Explanations: Justifying Decisions and Pointing to the Evidence*, Conference on Computer Vision and Pattern Recognition (June 18–22, 2018), https://arxiv.org/pdf/1802.08129.pdf.

[163] *See generally* MICHAEL BOHLANDER, PRINCIPLES OF GERMAN CRIMINAL PROCEDURE 145–57 (2012).

[164] The code does include an inquisitorial version of cross-examination, although in practice it is not used in any meaningful way. StPO § 239.

[165] *Cf. id.*§ 244, 250, 261.

nature of the process allows for information that would be considered hearsay in a common law system to be admitted.[166] Unwavering trust in professional judges' fact-finding capabilities and insight into human nature, as well as their purported impartiality, brought about statutory provisions conferring on the bench the power to provide only as much context as they see fit to witness or expert testimony.[167]

As the trier of fact, the bench is tasked with hearing all relevant evidence and assessing it freely (without bias) to determine its probative value.[168] An initial comparative look at German criminal procedure would reveal that it lacks credibility testing similar to that of adversarial jurisdictions.[169] However, the defense's right to question incriminating witnesses has been notably enhanced through the ECtHR's case law around confrontation rights found in Article 6 of the ECHR.[170] Notably, reference to the fair trial rights established under this provision appear to have gained importance with regard to machine evidence. In recent cases, these human rights seem to have served as a sort of secondary remedy for a defendant requesting knowledge parity where fact-finding is based on a digital evidentiary tool like a radar gun.[171] It reveals a gap in adequate remedies under traditional German law where a case file lacks relevant information and the prosecution or the bench are unwilling to vest the defense with the means to properly rebut a prosecutorial claim.[172]

The inquisitorial tradition's allowance of hearsay evidence in proceedings is one of several fundamental differences from the adversarial

[166] THOMAS WEIGEND, *Defense Rights in European Legal Systems under the Influence of the European Court of Human Rights*, *in* The Oxford Handbook of Criminal Process 165–188 (Darryl K. Brown, Jenia L. Turner & Bettina Weisser eds., 2019).

[167] StPO §244 ¶ 2 (The presiding judge is primarily responsible for deciding what evidence will be presented at the trial. The prosecution as well as the defense may propose additional pieces of evidence, but the court decides on the relevance and admissibility of the proposed evidence.).

[168] Damaška, *supra* note 14, at 446.

[169] Some defense lawyers have begun to point out the flaws of the inquisitorial system and test the boundaries of the tools available to them. *See, e.g.*, Ralf Neuhaus, *Kriminaltechnik für den Strafverteidiger – Eine Einführung in die Grundlagen*, 24 STRAFVERTEIDIGER-FORUM 393 (2006) (Ger.).

[170] Weigand, *supra* note 166, at 183.

[171] Wendt, *supra* note 122; Cierniak & Niehaus, *supra* note 98; *see also* Oberlandesgericht [OLG] Karlsruhe July 16, 2019, 1 Rb10 Ss 291/19 (Ger.); Verfassungsgerichtshof [VerfGH] Saarbrücken, Apr. 27, 2018, Lv 1/18; Kammergericht [KG], April 2, 2019, 3 Ws [B] 97/19 – 122 Ss 43/19 (Ger.).

[172] *Cf.* Benjamin Krenberger, Anmerkung zu Verfassungsgerichtshof (VerfGH) Saarbrücken, Apr. 27, 2018 Lv 1/18, 30 NEUE ZEITSCHRIFT FÜR VERKEHRSRECHT 282–83 (2018) (Ger.).

model.[173] At first glance, one might assume that the German system stands in stark contrast to the United States' insistence on a strict application of the hearsay rule.[174] However, a closer comparison of German and U.S. federal law would reveal some commonalities between the two systems. For instance, both take a similar approach with reports detailing laboratory findings at trial, including who must be called to the stand to discuss the conclusions.[175]

In Germany, mechanisms to account for trustworthiness in fact-finding might not be as visible as in adversarial jurisdictions. Like other inquisitorial systems, it relies heavily upon the various parties involved in a case to add to contribute to the establishment of facts throughout the investigation. This culminates in the bench summoning relevant witnesses to the oral hearing prior to the final fact-finding, always keeping in mind the principle of immediacy.[176]

This principle of immediacy is the most important safeguard in place in the German system and aims to achieve direct contact between the trier of fact and the source of information. Typically, the result would be that all witnesses must appear in court or, to quote from German procedural law, "[i]f the proof of a fact is based on the observation of a person, such person shall be examined at the main hearing. The examination shall not be replaced by reading out the record of a previous examination or reading out a written statement."[177] A problem arises in the case of digital evidentiary devices because their analysis is done outside the courtroom and behind the "closed doors" of the device.

Due to legal reforms and the significant degree of trust placed in the bench, an evaluation of the credibility of witness statements is left to the court. Witnesses are not assigned to a "side" (prosecution or defense), but instead are part of the court's overall truth-seeking mission.[178] The right of the prosecution and the defense to question witnesses is exercised in a manner less formal than the in-person cross-examination format of adversarial systems. Given this type of structure, calling for the functional equivalent of witness examination of a robot,

[173] Strictly speaking, there is no admissibility test in a German criminal proceeding.

[174] *See* Crawford v. Washington, 541 U.S. 36, 40 (2004).

[175] Bundesgerichtshof [BGH] [Federal Court of Justice] May 25, 2011, 2 StR, 2011, 585/10 (Ger.); Kathleen Schnoor, BEURTEILUNG DER SCHULDFÄHIGKEIT – EINE EMPIRISCHE UNTERSUCHUNG ZUM UMGANG DER JUSTIZ MIT SACHVERSTÄNDIGEN (2009) (Ger.).

[176] Although exceptions to the general rules, the prosecution and defense may bring their own witnesses to trial or ask the bench to summon a witness. *See* StPO § 244.

[177] *Id.* § 250.

[178] Namely, the "opinion rule," precluding conclusory factual statements by lay witnesses, or the "best evidence rule," requiring original documentation to prove the contents of a writing.

as a provider of quasi-testimonial statements, does not seem unreasonable in the German system.

That said, it is not clear how the German legal community would respond to such an idea, particularly given that over time the shield of immediacy has become porous. Although procedural law has generally been built upon the traditional dictum that a court must base its judgement solely on what is said or done during a public trial, German law today provides for a number of exceptions where depositions of absent witnesses can be admitted as evidence and, as a result, the original best evidence rule has lost ground.[179] Additionally, written records documenting the previous examination of witnesses, experts, or even a co-defendant can replace their oral examination where "illness, infirmity, or other insurmountable impediments" prevent their appearance at the main hearing "for a long or indefinite period," or where "the public prosecutor, defense counsel, and the accused agree to the reading out."[180]

This development poses both risks and potential solutions with regard to machine evidence. On the one hand, it provides space to thoroughly test AI-driven devices outside the courtroom, thereby providing relevant documentation for the trier of fact, and thus establishing the foundation for the assessment of credibility. On the other hand, it carries the risk that machine evidence will not be treated as testimonial evidence and could eventually enter the fact-finding domain without sufficient vetting. If we accept that AI conveys a message of its own through its evaluation of a human user's conduct, we would want to know the type of input perceived and how it led to a particular conclusion, similar to what would we would ask of a human witness on the stand. Therefore, the question remains as to how AI can be adequately evaluated to ensure trustworthiness.

b. United States

If federal courts regard machine evidence as a type of pre-recorded witness statement required to be cross-examined to ascertain the

[179] Damaška, *supra* note 14, at 425, 448 n.64 (drawing a line from Carl Mittermaier to English evidentiary law); *see also* Eser et al., *AE-Beweisaufnahme*, GA 2014, 1, 13 ff.; Thomas Weigend, *Das Konfrontationsrecht des Angeklagten – wesentliches Element eines fairen Verfahrens oder Fremdkörper im deutschen Strafprozess?*, *in* GESAMTE STRAFRECHTSWISSENSCHAFT IN INTERNATIONALER DIMENSION: FESTSCHRIFT FÜR JÜRGEN WOLTER ZUM 70. GEBURTSTAG AM 7. SEPTEMBER 2013, 1145 (Mark A. Zöller et al. eds., 2013) (Ger.).

[180] StPO § 251, para. 2.

veracity of its assertion, who would be called to the stand? The machine? The humans behind the machine?

Interestingly enough, this question has resulted in a fierce debate among U.S. legal scholars,[181] but it also hints at a crossroads between adversarial and inquisitorial systems. The hearsay exclusion is integral to the collective memory of the U.S. adversarial system. In *Crawford*, Justice Scalia explains why testimony before a magistrate is generally insufficient in a U.S. court despite being adequate in civil law jurisdictions.[182] While witness and expert depositions can be evaluated by both parties and the bench in the inquisitorial fact-finding process, such information remains inadmissible in adversarial proceedings unless they fall under an exception. This poses a problem where the reliability and credibility of evidence is based on complex technology that must be understood, tested, and adequately explained to the trier of fact. It is for this reason that several scholars have argued that the existing format of credibility testing should be broadened or that a similar deposition-style credibility test should be created for machine evidence.[183]

Christian Chessmann has put forward the idea of using impeachment techniques to bring the humans that created the relevant machine or software into the courtroom rather than trying to find equivalent means of putting a robot on the witness stand.[184] While a reasonable idea, it comes with its own problems for courts.[185] Given that cross-examining a witness often involves questioning of their credibility, a large number of people would have to be called into court in order to challenge data generated by an AI-driven device like an automated car because, at least in this case, it is dependent on a number of "driving assistants," each of which has its own source code and machine-learning standardization data.[186]

[181] *See, e.g.*, Roth, *supra* note 3, at 2046; Friedman, *supra* note 151, at 256–59.

[182] Crawford v. Washington, 541 U.S. 36, 53–56 (2004); *see also* People v. Lopez, 286 P.3d 469, 494 (Cal. 2012) (Liu, J., dissenting) ("[A]s a result of ever more powerful technologies, our justice system has increasingly relied on ex parte computerized determinations of critical facts in criminal proceedings—determinations once made by human beings. A crime lab's reliance on gas chromatography may be a marked improvement over less accurate or more subjective methods of determining blood-alcohol levels. The allure of such technology is its infallibility, its precision, its incorruptibility. But I wonder if that allure should prompt us to remain alert to constitutional concerns, lest we gradually recreate through machines instead of magistrates the civil law mode of ex parte production of evidence that constituted the 'principal evil at which the Confrontation Clause was directed.'") (quoting *Crawford*, 541 U.S. at 50.).

[183] Roth, *supra* note 30, at 1283–85, 1300–01.

[184] Chessman, *supra* note 5, at 220 n. 310.

[185] *Id.*

[186] Roth, *supra* note 30, at 1278.

As an early scholar in the field, Andrea Roth proposed a new path, arguing that the constant focus on hearsay, which by definition refers only to out-of-court statements by *people*, might be misleading in the digital age.[187] In a detailed analysis of the many potential flaws of intelligent machines, she encourages a new format for what would be the equivalent of impeaching a non-human entity. It would require diverging from typical courtroom testing but could potentially apply to both adversarial and inquisitorial systems.

Instead of an oral cross-examination of the programmer responsible for designing specific software, the machine's overall potential for faulty design and data production would be probed prior to trial by giving both sides access to certain information. Ideally, this would include the source code, standardization data, and training data (respecting trade secrets where applicable), as well as the possibility to experiment with the machine and analyze the algorithmic system.[188] Such testing of the machine could amount to an out-of-court "confrontation," which would need to be subsequently introduced to the court by an expert who participated in the evaluation.

5. Machine Evidence's Need for Translation Through Expert Testimony

Machine evidence and expert testimony are inextricably linked due to the fact that AI-generated data must be explained. In the German system, the bench, as the factfinder, typically appoints an expert when it feels it lacks the requisite expertise, but provides few tools for the defense to challenge court-appointed expert evidence. The U.S. system, with its partisan experts and one-sided presentations, risks the possibility that conclusions offered by forensic experts go beyond the boundaries of general scientific knowledge into the realm of misleading the trier of fact. This is particularly risky when indigent defendants are without the resources to hire respected and experienced experts.

As mentioned already, experts are crucial to the use of machine evidence in a criminal trial. Specialists must capture and clarify robot input data in addition to processing and assessing them for information relevant to fact-finding. Where an assessment or other data generated by AI is offered as evidence, it acts in some ways as a type of expert witness. For example, AI-driven systems' ability to detect patterns in facial

[187] Roth, *supra* note 3, at 2046.

[188] *See* Perel & Elkin-Koren, *supra* note 129, at 198–212; Roth, *supra* note 3, at 2050 (suggesting programmers could give live testimony before a type of scientific commission and return to this commission anytime the software is changed or updated).

movements is achieved through an ability to decipher patterns from a large pool of data that is beyond human comprehension. Such AI expertise exceeds human capacity, and human experts, even when given adequate time to question and evaluate AI, are often unable to explain all the details of the operational process and conclusions. This is the crux of the problem with machine data and it remains today something that must be addressed in a courtroom where humans require some type of explanation to help them evaluate the reliability and credibility of evidence.

When vetting human expert testimony, procedural laws generally allow for an evaluation of the factors contributing to an expert's finding and how they came to a certain conclusion. The functional equivalent of evaluating an expert testifying on behalf of AI-generated information is needed but would likely not involve the same explanations as in human testimony. Nevertheless, the trier of fact must understand the process of how the machine gathers information, evaluates it, and makes a determination. Such insight would ideally result from direct communication with the AI-driven device or software used but, as things stand today, we are limited to explanations by human experts, which are hampered by black box problems and the reality that most attempts to increase one's ability to explain AI processes will negatively affect the accuracy of the statement.[189] The question then becomes, what evidentiary safeguards can be put into place to address the potential risks to trustworthy fact-finding resulting from a human expert providing testimony for an AI-driven device that can neither speak for itself nor explain its assessment?

a. Germany

Allowing comprehensive out-of-court testing and the compilation of complex test results into a case file shared by all parties offers a useful approach for testing machine evidence, particularly where human experts are providing evidence about machine data. This format permits the sufficient development and evaluation of the strengths and weaknesses of an AI-driven device's ability to accurately assess human conduct. However, an inability to ensure impartiality and accurate self-assessment by the bench remains a weakness.[190]

Beyond testifying in court, a court-appointed expert can add to the trier of fact's knowledge through an in-depth out-of-court evaluation. Such an evaluation is then trusted by the court because the expert has

[189] *See* Doshi-Velez & Kortz, *supra* note 60.

[190] *See* Michael Bohlander, *supra* note 163, at 154–56, 170–71.

"been sworn generally to render opinions of the kind concerned."[191] The expert's report can be accepted by the bench as a written expertise under one of the exceptions to the principle of immediacy, but the expert may also be called into court to orally explain the report and be questioned to clarify any areas of doubt. Exceptionally, for instance where there is insufficient time to prepare a report in advance, a defense expert may come to court without having previously presented a written report.

Forgoing an oral report by an expert carries the inherent risk that important questions will not be asked. In the case of machine evidence, the "communication" with the software or device is entirely mediated by the expert without any scrutiny by the parties. Whether, and to what extent, a court will summon an expert to verbally explain their methods and findings depends on the value the bench places in listening to the expert in court (to support fact-finding) and the prosecution's and defense's opinions of a report.[192] The bench takes into account its obligation to seek the truth, based on the right to freely assess available evidence, and the necessity to rationalize and explain how facts were established the establishment of facts in a reasoned judgment to avoid reversal on appeal. This is especially so where the prosecution or defense has pointed out flaws in an expert opinion.[193] Where complex evidence is at issue, the inquisitorial process of building a case file provides sufficient opportunity to thoroughly monitor expertise. In the case of machine evidence, this would allow for investigation into the construction and operation of an AI-driven device, during which time parties could point out any potential flaws. This is, of course, dependent upon both sides having adequate resources and access to all the relevant information.

While fact-finding during the investigation and the preparation of a trial is shaped by the prosecution and the court, where the defense

[191] StPO § 256.

[192] In practice, when expert evidence is central to the charges, especially where the mental fitness of a defendant is in question, a judge will contact the prosecution and defense before appointing an expert to give them an opportunity to comment prior to summoning the expert to present her opinion during the oral hearing. *See, e.g.*, Klaus Detter,*Der Sachverständige im Strafverfahren - eine Bestandsaufnahme*, 18 NEUE ZEITSCHRIFT FÜR STRAFRECHT 57, 58 (1998) (Ger.). When expert evidence is more of a mechanical assessment, such as reading data from an automated vehicle, the prosecution typically contacts the expert during the investigatory stage to decide whether charges are to be brought at all and will add a written report to the case file. Generally, the choice of expert is accepted by the bench. *See* BeckOK StGB/Eschelbach [Beck's Online Commentary of German Penal Code] 45 StGB § 20, 100 (C.H. Beck 2020).

[193] Damaška, *supra* note 14, at 425, 454–55.

seeks to challenge the court's acceptance of a written expert opinion it has three options. First, a defendant can summon their own expert to appear at the oral hearing.[194] However, bringing forth and paying for one's own expert is the exception rather than the rule in an inquisitorial proceeding. Aside from the practical difficulties,[195] the defense must overcome the legal provision allowing a bench to refuse to hear an expert if, among other reasons,[196] it finds that the fact to be proven has already been established or it determines that the motion to hear the defense's expert has been made only for the purpose of delaying the proceedings.[197]

Second, the defense can make a motion to summon new expertise during preparation for the main hearing, at which time facts must be stated to support taking new evidence. This can be done after the prosecution has registered the case file with the court.[198] If the prosecution has commissioned an expert during the pretrial investigation, this expert is to be immediately available for the defense lawyer.[199]

The third and most powerful option is for a defendant to request that the trial court appoint an additional expert under Section 244.[200] The bench then must summon the expert unless it deems an additional expert opinion to be "superfluous because the matter is common knowledge, the fact to be proved is irrelevant to the decision or has already been proved, the evidence is wholly inappropriate or unobtainable, the application is made to protract the proceedings, or an important allegation which is intended to offer proof in exoneration of the defendant may be treated as if the alleged fact were true."[201] The court may also reject a motion for additional expert evidence "if the court itself possesses the necessary specialized knowledge" or "if the

[194] StPO § 220, para. 1.

[195] A person summoned by the defense must only comply if compensation is deposited with the court. *Id.* § 220, para. 2. Even where compensation is available, a defendant needs skilled counsel to question the expert. Notably, where the expert witness summoned by the defendant proves to be useful in clarifying the case, the court shall compensate him or her from the state treasury. *Id.* § 220, para. 3.

[196] Judges may not refuse an expert present at trial, summoned by the defendant, simply because they find that the testimony was unnecessary (because they are knowledgably themselves) or on the grounds that the defendant, not the bench, selected the expert. Obviously, the lines between legitimately and improperly refusing a defense expert are fluid, and where the court illegitimately refuses to hear such expert opinion, the final judgment may be appealed.

[197] *Id.* § 245, para. 2.

[198] *Id.* § 219, para. 1.

[199] *Id.* § 147, para. 3.

[200] *Id.* § 244, para. 3.

[201] *Id.*

opposite of the alleged fact has already been proved by the first expert opinion."[202] However, the rule does not apply "to cases where the professional competence of the first expert is in doubt, where his opinion is based upon incorrect factual suppositions, where the opinion contains contradictions, or where the new expert has means of research at his disposal which seem to be superior to the ones of an earlier expert."[203]

As the defense has access to the entire file it should, at least hypothetically, be aware of whether the expert opinion presented is in compliance with general standards or if new methods are available. Notably, the court's rejection of a motion for additional expert evidence can be appealed.[204] However, while there is vast case law addressing when the bench may reject such a motion, most do not delve into the issue of trustworthiness in fact-finding, which is so commonly questioned in adversarial proceedings, but instead focus on the issue of the bench's own proficiency in assessing a controversial question without the assistance of an expert.[205]

Unfortunately, where a bench believes that it is able to answer specific and complicated scientific questions of machine evidence itself (in what is often a less-than-realistic self-assessment), or where it is already inclined to believe a particular expert, neither party has a meaningful way to challenge such beliefs.[206] In those cases, as in others dealing with problems around the premature assessment of facts by benches, it may be especially difficult to ensure that the court is keeping an open mind about new types of evidence.[207]

In that respect, the German system has a blind spot in that it does not properly weigh the benefits of its nonpartisanship and generous amount of time granted to court-appointed experts against the risks of a lack of trustworthy fact-finding and open-minded benches. A 2015 expert law reform commission discusses an argument from judges that,

[202] *Id.* § 244, para. 4.

[203] *Id.*

[204] *Id.* § 337.

[205] Bundesgerichtshof [BGH] [Federal Court of Justice] July 6, 2011, 2 Strafrecht [StR] 124/11, 2011 (Ger.); Bundesgerichtshof [BGH] [Federal Court of Justice] July 9, 2015, 3 Strafrecht [StR] 516/14, 2015 (Ger.); Bundesgerichtshof [BGH] [Federal Court of Justice] July 12, 2017, 1 Strafrecht [StR] 408/16, 2017 (Ger.). See Vuille, *supra* note 100, at 231–32, for further details on the expert concept in Continental Europe; *see also* Jackson & Summers, *supra* note 94.

[206] Ulrich Eisenberg, BEWEISRECHT DER STPO: SPEZIALKOMMENTAR, 1518 (10th ed. 2017) (Ger.).

[207] Bundesgerichtshof [BGH] [Federal Court of Justice] May 30, 2000, 1 Strafrecht [StR] 582/99, 2000 (Ger.).

in an effort to work efficiently, "[i]n certain areas, the courts are reliant on continuous cooperation with experts whose expertise they are convinced of and who, in individual cases, also guarantee them rapid execution of short-term orders."[208]

b. United States

U.S. law offers a highly partisan structure to test expert evidence, means for thorough out-of-court testing (albeit underdeveloped), and a strong preference for the trier of fact to have unmediated access to anyone attempting to convey a specific opinion. Using the example of automated driving, this would translate into a desire to make visible any use of AI for evidentiary purposes in the courtroom in order to ensure trustworthiness in the justice process.[209] With respect to the assessment of machine evidence, the U.S. system is rather unyielding when it comes to out-of-court evidentiary testing, seemingly favoring more "direct" communication with AI technology. It also gives credence to the fact that despite attempts to improve the trustworthiness of fact-finding by providing objective scientific knowledge to the trier of fact, experts can be biased.

Although in the United States both parties select their own experts, judges are the gatekeepers of evidence, including expert opinions, by ruling on admissibility. The requirement that scientific and expert testimony in federal courts meet the *Daubert* standard of reliability provides the trier of fact with the context necessary to assess a source's credibility. For various reasons, including prevailing evidentiary law, judges are assigned the task of deciding the methodological issues of scientific evidence as a question of law under the four-tier test established in *Daubert v. Merrell Dow Pharm., Inc.*[210] and its progeny.[211] Judges must consider a

[208] The commission also advises judges and prosecutors to consult with the defense before selecting an expert. *See* Bundesministerium der Justiz und für Verbraucherschutz [Federal Ministry of Justice and Consumer Protection], Bericht der Expertenkommission zur effektiveren und praxistauglicheren Ausgestaltung des allgemeinen Strafverfahrens und des jugendgerichtlichen Verfahrens [Report of the Expert Commission on the Most Effective and Practicable Design of the General Criminal Procedure and the Juvenile Court Proceedings] 37 (2015), https://www.bmjv.de/SharedDocs/Downloads/DE/PDF/Abschlussbericht_Reform_StPO_Kommission.pdf (Ger.).

[209] *See* Daniel J. Capra, *Expanding (or Just Fixing) the Residual Exception to the Hearsay Rule*, 85 FORDHAM L. REV. 1577, 1581–4,1608–9 (2017).

[210] *See* Daubert v. Merrell Dow Pharm., Inc., 509 U.S. 579 (1993); Bernstein & Jackson, *supra* note 109; *see also* Mellon, *supra* note 63, for a detailed discussion.

[211] Gen. Elec. Co. v. Joiner, 522 U.S. 136, 145 (1997); Kumho Tire Co. v. Carmichael, 526 U.S. 137, 149 (1999).

number of factors in determining the admissibility of scientific and expert evidence, including: whether the scientific technique has successfully withstood testing; whether it has been subjected to peer review and publication; whether it has a known error rate and standards to control its operation; and whether it is generally accepted in a scientific community.[212]

If *Daubert* were used to determine the admissibility of machine evidence, it would help to exclude the most demonstrably unreliable machine evidence. However, the effectiveness of the *Daubert* test with regard to AI-generated data in human-robot interactions remains unclear as it is yet to be seen how consumer products that generate data will be offered as evidence. The question then becomes, how useful are such hearings beyond the most obvious cases? In all likelihood, most judges would admit machine evidence as long as some validation studies demonstrate that the machine's error rate is low and the methodology applied was sound,[213] but the existence of such studies remains to be seen given that commercial automated systems will almost certainly not be published and peer reviewed. It is clear that *Daubert* does not apply to machine-generated conclusions, but rather the witness statements in which they are included. Because of this, the question is really whether or not the *Daubert* test can and should be modified. Where an expert accompanying a machine into a courtroom is a "mere scrivener" for the machine's message,[214] the expert must pass the *Daubert* test; where a machine's assessment is used to prove a fact material to the case (like the assessment that a driver has in fact been sleepy) it must pass the *Daubert* test itself.

Overall, the effect of *Daubert* on the admissibility of machine evidence as a type of expert statement is unclear. Given that the government has the burden of proof in criminal cases, strict gatekeeping will directly impact new types of evidence and expertise that they offer. Where the defense takes a more active role, the central issue will likely become the evidence's relevance under Federal Rules of Evidence 401 and 403.[215] Indeed, getting the relevant information admitted into court and being able to thoroughly evaluate the workings of the AI employed are crucial.

Rule 17 of the Federal Rules of Criminal Procedure permits the defendant to subpoena evidence and witnesses independently, but still

[212] *See Daubert*, 509 U.S. at 593–94.

[213] See Roth, *supra* note 3, at 1981–82, for further discussion.

[214] *Id.* at 2032–33.

[215] Myers, *supra* note 71.

must comply with trade secret privileges. This rule could potentially be used to bring the defense's own expert evidence contesting the prosecution's case and also to force the government to give expanded discovery regarding the underlying source code, algorithms, and data for any forensic tools used.[216] Nevertheless, the judicial role as gatekeepers remains vital when the defense is the driving force in presenting novel scientific evidence. Courts appear to be reluctant to admit certain evidence, for instance if they think the defense is inclined to distort the science in hopes of creating reasonable doubt among the jurors, and rely on *Daubert* to keep unsound expert evidence out of the courtroom.[217] So, while in principle the defense has a means of calling expert evidence independently from the prosecution,[218] practically speaking, the situation (and the relevant case law) is more complicated.[219]

With regard to machine evidence generated by consumer products, Rule 17 is of special interest as it could be used as a tool to gain access to information held by third parties.[220] This may often be the case, for instance, when an AI-driven device in an automobile generated the data in question.[221] A Rule 17(c) subpoena is a traditional subpoena duces tecum for the production of items at trial. It also permits items to be "brought into court in advance . . . so that they may then be inspected in advance, for the purpose of course of enabling the party to

[216] Chessman, *supra* note 5, at 179.

[217] *See* David E. Bernstein & Eric G. Lasker, *Defending Daubert: It's Time to Amend Federal Rule of Evidence 702*, 57 WM. & MARY L. REV. 1, 5 (2015); *see also* Suedabeh Walker, *Drawing on Daubert: Bringing Reliability to the Forefront in the Admissibility of Eyewitness Identification Testimony*, 82 EMORY L. J. 1205, 1207 (2013).

[218] This is particularly the case when the government omits certain evidence from the prosecution but keeps it in a separate "investigatory file," does not consider potentially exonerating evidence, or when the defense wants to reference an item of evidence in the possession of a third party.

[219] *See* United States v. Armstrong, 621 F.2d 951, 954 (9th Cir. 1980); United States v. Nixon, 418 U.S. 683 (1974); Pennsylvania v. Ritchie, 480 U.S. 39 (1987); United States v. Soape, 169 F.3d 257, 269 (5th Cir. 1999); Miller, *supra* note 136, at 326.

[220] *See, e.g.*, United States v. Yee, 129 F.R.D. 629 (N.D. Ohio 1990) (finding that laboratory matching criteria and standards, environmental insult tests, population data, and proficiency testing data held by the FBI are discoverable based on Rule 16(a)(1)(C)). *But see* United States v. Iglesias, 881 F.2d 1519, 1524 (9th Cir. 1989) (finding that log notes produced in the testing process are not discoverable as scientific reports).

[221] FED. R. CRIM. P. 17(c)(1) ("A subpoena may order the witness to produce any books, papers, documents, data, or other objects the subpoena designates. The court may direct the witness to produce the designated items in court before trial or before they are to be offered in evidence. When the items arrive, the court may permit the parties and their attorneys to inspect all or part of them.").

see whether he can use it or whether he wants to use it."[222] According to Rule 17 (b), the court must decide whether the public will pay the fee for such subpoenas where a defendant can show his or her inability to pay. Courts tend to be hesitant to do so where it is suspected that the defendant seeks to interpret the science in such a way that it creates reasonable doubt. The motion is then denied, at least in part, to prevent unnecessary public expenses, but also due to a fear of allowing "hocus-pocus" into the courtroom.[223]

The case law involving the use of digital tools in forensic settings, like DNA testing, suggests something akin to the evidentiary life cycle described above.[224] Today, the trend is toward a stricter standard, and the Advisory Committee on the Federal Rules of Evidence has discussed the potential addition of Rule 707, which would limit judicial discretion in an effort to enhance the reliability of expert evidence.[225] It would also provide greater inclusion of out-of-court statements from which courts expect the trier of fact to gain a better understanding of a complex evidentiary issue.

6. New Mechanisms of Contextualization and Credibility Testing

There is a need for new mechanisms to not only contextualize and test the credibility of machine evidence, but also to enable the trier of fact to assess the reliability of machine evidence. In the United States, this could be achieved by a partial separation of credibility testing from the prevailing courtroom-centered hearsay model, thereby initiating a new approach to confrontation rights in the digital age, including the right to evaluate AI-driven devices out of court.[226] By contrast, in

[222] Bowman Dairy Co. v. United States, 341 U.S. 214, 222 n.5 (1951).

[223] *See* Saltzburg, *supra* note 149, at 1720.

[224] *See supra* Part III.B.3.

[225] Memorandum from Daniel J. Capra, Reporter, Advisory Comm. on Evidence Rules, to Advisory Comm. on Evidence Rules, Symposium on Forensic Expert Testimony, Daubert and Rule 702 (Oct. 1, 2017), *in* ADVISORY COMMITTEE ON RULES OF EVIDENCE OCTOBER 2017 AGENDA BOOK 371, 381 (2017) (alterations in original) (proposing a new Federal Rule of Evidence Rule 707), http://www.uscourts.gov/sites/default/files/a3_0.pdf ("If a witness is testifying on the basis of a forensic examination [conducted to determine whether an evidentiary sample is similar or identical to a source sample], [or: "testifying to a forensic identification"] the proponent must prove the following in addition to satisfying the requirements of Rule 702: (a) the witness's method is repeatable, reproducible, and accurate — as shown by empirical studies conducted under conditions appropriate to its intended use; (b) the witness is capable of applying the method reliably. . . and actually did so; and (c) the witness accurately states. . . the probative value of [the meaning of] any similarity or match between the evidentiary sample and the source sample."); *see also* Saltzburg, *supra* note 149, at 1709.

[226] Roth, *supra* note 3, at 2048.

Germany, new tools must be provided to the defense to allow a bench's fact-finding and expert evidence to be meaningfully challenged; it is particularly important that the defense is granted access to any data that is at the disposal of the court-appointed expert.

A comparative analysis of the inquisitorial and adversarial criminal justice systems revealed that there are new evidentiary problems should machine evidence enter the courtroom and that there may not be one single solution. An increase in human-robot interaction will unlock the potential to access large quantities of information provided we are willing to overlook the significant questions around its reliability and our limited means of adequately evaluating such information. Additionally, the use of such data raises further concerns regarding trade secrecy and privacy, but these issues are beyond the scope of this Article.[227]

Nevertheless, the contrasting analysis provides valuable information about the distinct nature of machine evidence and a clearer picture of the evidentiary problems AI may cause in courtrooms across jurisdictions. Traditionally, both the inquisitorial and adversarial systems have addressed difficulties in maintaining trustworthy fact-finding by pointing out human errors in the evidentiary procedure. Both systems will have to modify their approach if they seek to introduce machine evidence into criminal proceedings. In some ways, the conclusion is the same in both jurisdictions: AI's unique status must be acknowledged and the message it conveys needs to be made visible to the parties, the court, and the public.

How AI is best made visible in the courtroom is strongly linked to the nuances of each criminal justice system. In the United States, a system proffered by Andrea Roth[228] proposes pre-trial credibility testing of the front-end design, input, and operation protocols of machine evidence. Such meaningful access to the device in question before trial could allow for a partisan review of the machine's functioning. Eventually, valid contextualization and credibility testing will depend on the machine or software's design and construction. Beyond domestic singularities and varying technology, a normative approach to ensure reliable fact-finding still must be determined. In doing so, the two systems can increase trustworthiness in fact-finding when machine evidence is at issue by departing from prevailing methods of credibility testing.

[227] *See* Nawara*, supra* note 105, at 614 (for issues around trade secrets); *see also* Wexler, *supra* note 78; Ram, *supra* note 78, at 665–683, 701–03.

[228] Roth, *supra* note 3, at 2028.

In an adversarial system, this would mean changes to the courtroom-centered model of testimony.[229] To ensure the credibility of an intelligent machine and the reliability of its evidence, the complexity of the underlying technology might be better scrutinized outside the courtroom independent from a case.[230] Experts evaluating the machine's design, learning pattern, source code, and other programming might be better able to make an initial decision about its reliability and credibility outside the courtroom, which could act as an individual assessment of evidence offered in a case.[231] These types of evidentiary proxies could be accepted as necessary contributions to an evaluation of trustworthiness under the circumstances.[232]

In an inquisitorial system, features within the fact-finding procedure that result in unyielding trust in traditional defense tools, which are based upon the presumption that a bench is able to maintain an open mind throughout the proceeding, must be addressed. This could be accomplished by granting a right (where necessary) to challenge both judges and experts that goes beyond the current framework of simply allowing a defendant to prepare his or her defense using the reports in the case file. This is because the available evidentiary tools today may fail to address the crucial issues around the reliability of machine evidence.[233] With regard to machine evidence, the approach would need to shift from one that is bench-dominated toward one that is more examination-oriented whereby all incriminating witnesses are evaluated pursuant to Article 6 of the ECHR.[234] This move toward a more adversarial method of dealing with machine evidence has been adopted by a majority of the Higher Regional Courts in Germany and shows promise as a means of vetting digital evidentiary tools.[235]

[229] *See, e.g.*, Celentino, *supra* note 45, at 1331–33.

[230] Roth, *supra* note 3, at 2028.

[231] *See* Perel & Elkin-Koren, *supra* note 129, at 186–98.

[232] *Cf.* Damaška, *supra* note 14, at 425, 446.

[233] *See* Jürgen Cierniak, *Prozessuale Anforderungen an den Nachweis von Verkehrsverstössen*, 12 ZEITSCHRIFT FÜR SCHADENSRECHT 664 (2012) (Ger.).

[234] European Convention for the Protection of Human Rights and Fundamental Freedoms art. 6, Sept. 3, 1953, 213 U.N.T.S. 222; *see also* Vuille, Lupària & Taroni, *supra* note 100, at 59–62; Cierniak & Niehaus, *supra* note 98.

[235] Oberlandesgericht [OLG] Bamberg, June 13, 2018, 3 Ss Owi 626/18 (Ger.); Oberlandesgericht [OLG] Karlsruhe July 16, 2019, 1 Rb 10 Ss 291/19 (Ger.); Verfassungsgerichtshof [VERFGH] Saarbrücken, Apr. 27, 2018 Lv 1/18; Kammergericht [KG] April 2, 2019, 3 Ws [B] 97/19 – 122 Ss 43/19 (Ger.).

IV. Conclusion

Where machine evidence is proffered as evidence in a criminal trial, it must be adequately contextualized and tested for reliability. Such evidence—just like human testimony—is not infallible.[236] Especially where the digital output of an opaque device, initially produced as technology for a consumer need, is accepted as a conduit of fact or circumstantial evidence, legislatures and courts must address this issue both open-mindedly and critically.[237]

With the rise of AI in all areas of human life, it seems especially important that legal scholarship point out that our presumptions of regularity and impartiality in machine workings are often inaccurate.[238] Many factors have to be taken into account. Research conducted across a variety of fields has demonstrated that the physical shape and cognitive capacity of an AI-driven device directly affects our perception of its reliability, soundness, and overall "character."[239] We seldom realize this fact and even more rarely question what features of any device should trigger confidence or doubt in the reliability of is findings.[240]

From a legal perspective, the use of AI-generated evidence in criminal proceedings remains an enigma. Laws do not provide rules around scrutinizing "intelligent machines" for credibility and, as long as

[236] *See* Roth, *supra* note 30, at 1283–85, 1300–01; Erin Murphy, *Databases, Doctrine & Constitutional Criminal Procedure*, 37 FORDHAM URB. L.J. 803, 825–26 (2010); Ellen M. Ayoob, Aaron Steinfeld & Richard Grace, *Identification of an "Appropriate" Drowsy Driver Detection Interface for Commercial Vehicle Operations*, 47 PROC. HUM. FACTORS & ERGONOMICS SOC'Y ANN. MEETING 1835, 1840–44 (2003).

[237] *See* Sergey Bratus, Ashlyn Lembree & Anna Shubina, *Software on the Witness Stand: What Should It Take for Us to Trust It?*, *in* TRUST AND TRUSTWORTHY COMPUTING 396–416 (Alessandro Acquisti, Sean W. Smith & Ahmad-Reza Sadehi eds., 2010) (identifying and critiquing several criminal cases); Eric Van Buskirk & Vincent T. Liu, *Digital Evidence: Challenging the Presumption of Reliability*, 1 J. DIGITAL FORENSIC PRAC. 19, 20–21 (2006).

[238] Murphy, *supra* note 236, at 804; Garrett, *supra* note 5, at 213; Chessman, *supra* note 5; Roth, *supra* note 3, at 1989–99.

[239] See, for example, Kate Darling, Palash Nandy & Cynthia Breazeal, *Empathic Concern and the Effect of Stories in Human-Robot Interaction*, PROC. 24TH IEEE INT'L SYMP. ON ROBOT HUM. INTERACTIVE COMM. 772–75 (2015), for a detailed discussion; *see also* Jacqueline M. Kory et al., *Effects of Framing a Robot as a Social Agent or as a Machine on Children's Social Behavior*, Proc. 25th IEEE INT'L SYMP. ON ROBOT HUM. INTERACTIVE COMM. 689–93 (2016).

[240] While front-design and algorithms will certainly be significant, other elements might also prove to be important. Were a driving assistant to have an anthropomorphic element *and* the ability to provide an account of its actions in court (e.g., "At 11pm, I advised the driver to take a break, but she overruled my advice."), the demand for credibility-testing might be stronger. Whether the pitch, volume, tone, and accent of the machine's voice will impact its perceived reliability will also likely become an issue.

machines and software lack the necessary characteristics to be meaningfully cross-examined about their conclusions, the law will continue to treat them as if they do not convey a message of their own and are nothing more than number-crunching tools. Such a cursory approach has been criticized as falling short in exposing all the potential risks to reliable fact-finding in criminal proceedings.[241]

A comparative perspective helps to better understand how the issue of machine evidence could potentially be resolved. The adversarial system has created tools to scrutinize evidence for reliability, while the inquisitorial system has developed a model of successive out-of-court evidence gathering that allows for all parties to meaningfully evaluate complex evidence.

The basic problem of machine evidence is the same for all jurisdictions. How can data generated by AI be adequately inspected given the fact that it cannot be vetted like human witnesses, yet still might contribute human-like biases through the way it processes data and assesses situations? How do we utilize such data knowing that machines and software, while not visible in a courtroom, are potentially active players in the events leading to a prosecution?

How to approach a new issue, like the use of machine evidence in criminal proceedings, is a dilemma well-known to other areas of the law. [242] Do we begin from a technical or legal standpoint, i.e., code for the law or law for the code?[243] The decision at this stage will not only determine whether adversarial and inquisitorial systems face similar or different problems going forward, but also whether or not fact-finding will remain the familiar human-centered procedure we know, which is focused on providing transparent and (ideally) objective information to the trier of fact to aid in decision-making.

Artificial Intelligence designed to interact with humans to meet a consumer need comes with embedded values.[244] In general, its formation does not align with the relevant legal norms on evidence law, and certainly does not comport to the typical fact-finding procedure or constitutional guarantees of a criminal trial. To believe that human parties to a criminal trial will be willing and able to decode the digital

[241] Mellon, *supra* note 63, at 1101; Pamela S. Katz, *Expert Robot: Using Artificial Intelligence to Assist Judges in Admitting Scientific Expert Testimony*, 24 ALB. L.J. SCI. & TECH. 1, 36 (2014).

[242] *See* Hildebrandt, *supra* note 2, at 165.

[243] Roger Brownsword, *What the World Needs Now: Techno-Regulation, Human Rights and Human Dignity*, *in* GLOBAL GOVERNANCE AND THE QUEST FOR JUSTICE 203 (Roger Brownsword ed., 2004); Ronald Leenes, *Framing Techno-Regulation: An Exploration of State and Non-State Regulation by Technology*, 5 LEGISPRUDENCE 143 (2011).

[244] Hildebrandt, *supra* note 2, at 166.

infrastructure in pursuit of the truth without having access to new evidentiary tools is naïve.

Therefore, now is the time to prepare for fact-finding in ambient intelligent environments. To do so, we first must understand the characteristics of the various types of machine evidence and work with qualified experts to both understand the technology and explain the underlying legal concepts.[245] Only then can AI-driven devices be meaningfully evaluated in the courtroom from all necessary angles. If one addresses the problem of integrating machine evidence into the establishment of facts in criminal trials, from a technical standpoint (thereby requiring technology to serve the law), the adversarial and the inquisitorial systems would face the same, albeit monumental, challenges. Both could theoretically take similar action, like certifying AI-driven devices, providing open access to source code, and specifying machine learning parameters. If new legal solutions were to be pursued, each system would need to find its own answer, but could still learn from the other. It is likely that machine evidence will follow the life cycle of technological evidence: Initially deemed too new to be reliable; then new but subject to testing and as a result, regarded as generally reliable; and finally, it may reach the point of being blindly trusted. A look back into the recent past of criminal justice teaches us that reversing the evidentiary cycle is an uphill battle and one that is preceded by a great deal of human suffering because of judicial errors. It will be important to understand how procedural safeguards apply to AI in the courtroom.

Regardless of whether AI becomes a new tool to convict[246] or acquit,[247] we must ensure trustworthiness in the fact-finding process where machine evidence is used in criminal proceedings. In general, humans trust each other's testimony despite a great deal of evidence questioning its reliability. Assumedly, we find it convincing because we can relate to human perception and experience; in a word, we possess empathy. Machine evidence could attempt to create a similar impression and perhaps what they lack in human characteristics they make up for with purported objectivity. It is not entirely clear why we humans are wired to believe the statements of our fellow human beings. Perhaps it is because we trust in the inherent goodness of people or we assume that the fear of punishment for perjury will prevent them from

[245] *See* Perel & Elkin-Koren, *supra* note 129, at 185.

[246] Historically, new "objective" technology was first used against defendants; more recently defendants have been using technology to their advantage. Roth, *supra* note 30, at 1254–64.

[247] Fairfield & Luna, *supra* note 49, at 990.

lying. However, AI, as we know it today, is subject to none of these constraints.

These are issues which must be urgently addressed if the law is to keep up with the rapid pace of advancing technology and are best solved through mutual learning between adversarial and inquisitorial justice systems. No evidentiary system is perfect, but the U.S. system prides itself on its strength, flexibility, and willingness to experiment with new approaches.[248] This is the opportunity to showcase such virtues. The trier of fact ought not to be faced with a new reality that limits reliability and credibility testing just because a new and unfamiliar type of evidence is built upon a particular design that cannot be meaningfully challenged in today's courtrooms. In order to preserve the authenticity and legitimacy of fact-finding in a criminal trial, it must remain human-centered.

[248] Mirjan R. Damaška, EVIDENCE LAW ADRIFT 151 (1997).