Tulsa Law Review

Volume 59 | Number 1

Winter 2024

AI-Based Evidence in Criminal Trials?

Sabine Gless

Frederic Lederer

Thomas Weigend

AI-BASED EVIDENCE IN CRIMINAL TRIALS?

Sabine Gless, Fredric Lederer, & Thomas Weigend[1]

---

1. Sabine Gless is Professor of Criminal Law and Criminal Procedure, University of Basel (Switzerland); Fredric Lederer is Chancellor Professor of Law and Director, Center for Legal & Court Technology, William & Mary Law School; Thomas Weigend is Professor Emeritus of Criminal Law, University of Cologne (Germany). The authors appreciate the support of colleagues, especially Jeffrey Bellin, Cabell Research Professor and Mill E. Godwin Jr. Professor of Law, and Tracy Byrd, Center for Legal & Court Technology Administrator, both of William & Mary Law School, Xuan Sharon Di, Associate Professor at Columbia University's Engineering Department, Erin Murphy, Norman Dorsen Professor of Civil Liberties, and Stephen Schulhofer, Robert B. McKay Professor of Law, both of New York University School of Law, whose assistance in critiquing this article is deeply appreciated. The authors also thank Emily Silverman of the Max Planck Institute for the Study of Crime, Security and Law (Freiburg, Germany) for providing inspiring inputs.

Smart devices are increasingly the origin of critical criminal case data. The importance of such data, especially data generated when using modern automobiles, is likely to become even more important as increasingly complex methods of machine learning lead to AI-based evidence being autonomously generated by devices. This article reviews the admissibility of such evidence from both American and German perspectives. As a result of this comparative approach, the authors conclude that American evidence law could be improved by borrowing aspects of the expert testimony approaches used in Germany's "inquisitorial" court system.

I. INTRODUCTION

A. Setting the Stage

We call BMW 7500i, Vehicle Number 12778899, to testify:

Q: Where were you at 8:42 p.m. on February 28th?

A: According to my navigation system and the stored data from it, we were located just past the intersection of Max Planck Drive and Rose Street.

Q: What occurred then?

A: My forward sensors detected that the automobile in front of us was slowing; I sounded the driver audio and video collision warning.

Q: What happened next?

A: My driver ignored the warning, accelerated, and attempted to pass the automobile to its right. He failed to do so, and we hit the left rear of the car.

Q: What proof do you have of this other than the conclusion you just gave?

A: In addition to the raw data in my storage nodes, I have a digital audio-video recorder that is automatically turned on when the collision alert is live; I can show you that recording now.

Is this science fiction or fantasy? In one sense, this is very real as our data is collected and analyzed through the technology we use on a daily basis. Although we are not aware of a case in which an automobile has "testified" as described above, it is not unreasonable to think that it might occur in the future. The automobile example above is loosely related to an actual case.[2] In 2016, the Swiss news media reported that the driver

2. *Swiss Politician Fined Over Crash That Injured 17-Year-Old*, THE LOCAL (Oct. 31, 2016), *https://www.thelocal.ch/20161031/swiss-politician-fined-over-crash-that-injured-17-year-old.*

assistance system embedded in a car had previously (and repeatedly) alerted the driver to driving errors related to fatigue and were ignored.[3] Drowsiness detection systems use raw data from a car's sensors, including data points such as lane departures, but also information about a driver's steering, body tension, seat position, and eyelid movements.[4] With the help of complex algorithms, the system constantly evaluates this data for signs of drowsiness.[5] If the system determines that the drowsiness threshold established by the programmer has been met, it will issue an alert to the driver, record the warning in its system, and possibly intervene by taking over the steering function.[6] In the Swiss case, the driver was accused of being unfit to drive due to drowsiness, causing the victim's injuries.[7] This was based upon data generated by his car's drowsiness detection system, which produced what Andrea Roth has coined "machine conveyance," or functionally, a form of incriminating statements.[8] The driver, a politician, was accused of causing the victim's injury by driving despite being unfit to do so because of drowsiness.[9] Ultimately, the driver accepted a summary penalty order[10] for the offense of "causing bodily harm through negligent driving."[11]

Although we will return to our futuristic BMW example, the focus of this article is on devices that collect, store, and "interpret" data, and especially the enhanced data storage and safety systems of today's automobiles.[12] In fact, they collect and store far more

3. *Id.*

4. Muhammad Ramzan, et al., *A Survey on State-of-the-Art Drowsiness Detection Techniques*, 7 IEEE ACCESS, 61904, 61906–07 (2019).

5. *Id.* at 61909–10, 61916–17.

6. *Id*. at 61904, 61909–10, 61914.

7. *Swiss Politician Fined Over Crash That Injured 17-Year-Old*, THE LOCAL (Oct. 31, 2016), *https://www.thelocal.ch/20161031/swiss-politician-fined-over-crash-that-injured-17-year-old. But see* Winnie Hu & Nate Schweber, *Bus Driver Found Not Guilty of Manslaughter in I-95 Crash,* N.Y. TIMES (Dec 7, 2012), https://www.nytimes.com/2012/12/08/nyregion/ophadell-williams-driver-in-fatal-bus-crash-found-not-guilty-of-manslaughter.html (finding deadly crash in New York City highlighted the difficulty of prosecuting accidents involving drowsy drivers).

8. Andrea Roth, *Machine Testimony*, 126 YALE L.J. 1972, 1976 n.11 (2017).

9. *Swiss Politician Fined Over Crash That Injured 17-Year-Old*, THE LOCAL (Oct. 31, 2016), *https://www.thelocal.ch/20161031/swiss-politician-fined-over-crash-that-injured-17-year-old.*

10. In Switzerland, the public prosecutor's office issues a penalty order (e.g., fine, custodial sentence) for criminal offenses where responsibility has been adequately established and the accused has not filed a rejection within ten days. SCHWEIZERISCHES STRAFGESETZBUCH [STGB], CODE PÉNAL SUISSE [CP], CODICE PENALE SVIZZERO [CP] [CRIMINAL CODE] Oct. 5, 2007, SR 312, RS 312, art. 352 (Switz.).

11. *Swiss Politician Fined Over Crash That Injured 17-Year-Old*, THE LOCAL (Oct. 31, 2016), https://www.thelocal.ch/20161031/swiss-politician-fined-over-crash-that-injured-17-year-old.

12. New cars registered in the EU must be equipped with various driver assistance systems designed to enhance road safety, such as emergency lane-keeping systems and driver fatigue warning systems, 2019 O.J. (L 325) 10–15. Safeguards and updated rules for the approval of motor vehicles with such technology have also been implemented, 2018 O.J. (L 151) 1, 2.

data than one might expect.[13] In addition to data on speed[14] and braking,[15] cars record "Black Box" data (meaning that we cannot understand how the device reached its conclusion), including alerts issued by safety devices such as drowsiness detection systems, or certain infotainment data.[16] Private companies offer sophisticated products that can access the infotainment systems of many newer vehicles.[17]

The Berla device ("Berla"),[18] for instance, enables the police to ascertain navigation data, which tells the police where a car has been driven at any given time. More concerning, if a driver connects his or her cell phone to the vehicle, the Berla gives police access to the cell phone data transferred to the car during the time the phone was connected to the vehicle.[19] It's not just automobiles that collect data: Fitness trackers and smart phones also gather a host of information about their users.[20] Amazon's Alexa and similar home smart devices similarly collect data and raise issues about privacy.[21] In looking at existing and foreseeable developments of AI, we can say that Isaac Asimov's self-aware intelligent robots[22] do not yet exist, but the data from what may be their progenitors surely does.

B. Device Evidence

In this article, we address the question of how and under what conditions data generated by information technology ("IT") devices, which we will call "device evidence," may be admitted in criminal trials. Devices referenced in this article are those governed by firmware and software, i.e., instructions in the form of computer code that ordinarily are fixed and that can be examined.[23] The special feature of device evidence is its autonomous

---

13. *See, e.g.,* Nhien-An Le-Khac et al., *Smart Vehicle Forensics: Challenges and Case Study*, 109 FUTURE GENERATION COMPUTER SYS. 500, 500, 508 (2020); Ryan Calo, *Robotics and the Lessons of Cyberlaw*, 103 CAL. L. REV. 513, 536 (2015).

14. A tachometer, for example, collects information on the four wheels' rotational speed from sensors at each wheel and employs an algorithm for calculating the average. DEWESoft, *Angle Measurement* 1, 11 (2023), https://training.dewesoft.com/storage/pro/courses/angle-measurement.pdf. The system can also calculate differences among the four separate wheel speeds. *Id.* at 63. In the event of an accident, this calculation can subsequently help an expert in assessing the status of the vehicle on the road, determining whether the vehicle was still controlled by the driver. *See id.* Modern automobiles store the data in both individual wheel form and a composite speed form, both of which can be extracted. *Id.* at 11.

15. For information on the recording of such data (like vehicle speed, throttle position, and brake activity), *see generally* SOC'Y OF AUTOMOBILE ENGINEERS [SAE], *Surface Vehicle Recommended Practice, Event Data Recorder*, J1698 (May 2014), https://www.sae.org/standards/content/j1698_201405/.

16. *Cf.* Minyoung Kim et al., *Implementation of smart car infotainment system including black box and self-diagnosis function*, 8 INT'L J. OF SOFTWARE ENG'G & ITS APPLICATIONS 267, 273–74 (2014). "Infotainment systems" have also been defined as "in-vehicle information systems (IVIS)." DAVID L. STRAYER ET AL., VISUAL AND COGNITIVE DEMANDS OF USING APPLE CARPLAY, GOOGLE'S ANDROID AUTO AND FIVE DIFFERENT OEM INFOTAINMENT SYSTEMS, AAA FOUNDATION FOR TRAFFIC SAFETY 36 (2018).

17. *See Discover Vehicle Forensics*, BERLA, https://berla.co/discover/ (last visited Oct. 11, 2023).

18. *See Discover Vehicle Forensics*, BERLA, https://berla.co/discover/ (last visited Oct. 11, 2023).

19. Adam M. Gershowitz, *The Tesla Meets the Fourth Amendment,* BYU L. REV. 1135, 1139 (2023).

20. Alexis Rodis, *Fitbit Data and the Fourth Amendment: Why the Collection of Data from a Fitbit Constitutes a Search and Should Require a Warrant in Light of Carpenter v. United States,* 29 WM. & MARY BILL RTS. J. 533, 535 (2020).

21. Lauren Chlouber Howell, *Alexa Hears with Her Little Ears – But Does She Have the Privilege?*, 52 ST. MARY'S L.J. 837, 843 (2021).

22. *See generally* ISAAC ASIMOV, I ROBOT (1950).

23. *What is Firmware?*, ALWAREBYTES, https://www.malwarebytes.com/cybersecurity/computer/what-is-firmware (last visited Oct. 18, 2023). Subject to possible legal restrictions such as a trade secret privilege, *cf.*

production: There is no human being that controls the production of the evidence, but rather, the device itself produces the data in accordance with its program.[24] Devices are typically separable or inseparable parts of a physical object, such as the hard disc of a computer or the data storage system in a car.[25] The core question is whether and how a proponent of device evidence can establish the accuracy of the data produced and its reliability (i.e., the degree to which a result can be expected to occur again under equal circumstances).[26] Devices differ with regard to their functions. Some devices, such as breathalyzers or radar guns, are designed for forensic purposes.[27] Others are consumer products such as cars, smart watches, fitness trackers, and medical devices such as pacemakers.[28] If data generated by the latter type of devices is used as forensic evidence, "function creep"[29] can occur. This expression refers to a situation where a device designed for a specific purpose is used for a different purpose for which it has not been fully evaluated.[30] For example, a drowsiness monitoring system installed in a car to increase traffic safety would acquire a new function if its data was used against the driver as evidence in a criminal court. This "function creep" can unfairly disadvantage the driver because car producers may wish to reduce their own potential liability by calibrating the system in a way that triggers an alarm at the very first sign of potential drowsiness.[31]

Devices also differ with respect to their sophistication. Some are limited to collecting and storing data (Type 1) and typically function in a rule-based way, producing data according to their fixed coding. Other devices draw conclusions from the data they collect and act upon their conclusions, having been trained to solve specific tasks using algorithms and statistical models (Type 2).[32] They apply their findings to new situations

---

Rebecca Wexler, *Life, Liberty, and Trade Secrets: Intellectual Property in the Criminal Justice System*, 70 STAN. L. REV. 1343, 1349–50 (2018); Roth, *supra* note 8, at 2028.

24. *What is IoT?*, ORACLE, https://www.oracle.com/internet-of-things/what-is-iot/#why-is-iot-important (last visited Oct. 18, 2023).

25. *See e.g.*, Teresa Reidt, *What is Firmware and What Does it Do?,* EMTERIA (Feb. 17, 2022, 7:30 AM), https://emteria.com/learn/firmware.

26. *See* Samuel R. Gross & Jennifer L. Mnookin, *Expert Information and Expert Evidence: A Preliminary Taxonomy*, 34 SETON HALL L. REV. 141, 143–44 (2003).

27. *History of the Breathalyzer, Invented by Robert F. Borkenstein*, THE WILSON LAW FIRM, https://www.tkevinwilsonlawyer.com/library/history-of-the-breathalyzer.cfm#:~:text=The%20Breathalyzer%20gave%20law%20enforcement,courts%20of%20law%20as%20evidence (last visited Oct. 18, 2023); *70+ Years in the Making: Inside the Incredible History of the Police Speed* Gun, KUSTOM SIGNALS, INC., https://kustomsignals.com/blog/100-plus-years-in-the-making-the-incredible-history-of-the-police-speed-gun (last visited Oct. 28, 2023).

28. Anthony Corbo, *What is Consumer Technology?*, BUILT IN: CONSUMER TECHNOLOGY (Oct. 18, 2022), https://builtin.com/consumer-tech#. The given use of a device may be important as a medical device that collects highly personal health data, for example, may raise privacy and other concerns distinct from a device that collects and stores weather data. Jason Peres da Silva, *Privacy Data Ethics of Wearable Digital Health Technology*, The Warren Alpert Medical School: Center for Digital Health (May 4, 2023), https://digitalhealth.med.brown.edu/news/2023-05-04/ethics-wearables. However, for the purposes of this article we will address only the device's function to record and evaluate data, and then potentially to act on its conclusions.

29. Paul W. Grimm et al., *Artificial Intelligence as Evidence*, 19 NW. J. TECH. & INTELL. PROP. 9, 51 (2021).

30. *Id.*

31. For methods of training and adjusting drowsiness detection systems, *see generally* Elena Magán et al., *Driver Drowsiness Detection by Applying Deep Learning Techniques to Sequences of Images*, 12 APPLIED. SCIS. 1145 (2022); Bakheet Samy & Al-Hamadi Ayoub, *A Framework for Instantaneous Driver Drowsiness Detection Based on Improved HOG Features and Naïve Bayesian Classification*, 11 BRAIN SCIS. 240 (2021).

32. *See generally* Patrick Grieve, *Deep Learning vs. Machine Learning*, ZENDESK BLOG (last updated Sept. 20, 2023), https://www.zendesk.com/blog/machine-learning-and-deep-learning/.

without following explicit instructions.[33] The hallmark of Type 2 devices is a digital layer of "intelligence" added through machine learning so that these devices can evaluate data and determine whether and when to act on these evaluations.[34] There are many Type 2 devices that can be employed as evidence in court; for example, an automobile's driver assistance system or facial recognition software which can be used to compare images and determine the identity of persons recorded.[35] Such devices can be designed especially for government purposes or for consumer needs.[36] Other examples of Type 2 devices are software that tracks the location of certain persons or objects, such as fitness trackers, Google Earth or GPS devices,[37] and smart robot vacuum cleaners that "identify" (and avoid) obstacles like toys, pet waste, or cords.[38] Even more sophisticated devices can modify their operations based on their experience (Type 3).[39] Examples are smart grids and autonomously driving cars, and facial recognition devices ranging from smart door bells to sophisticated identification systems, which can not only adapt but also optimize their own code.[40] Such devices aim for greater efficiency than the devices currently in use.[41] If a car can learn on the street and independently adapt its safety features to individual drivers, it could potentially achieve greater safety.[42] For such self-optimizing Type 3 devices, we use the term *Artificial Intelligence (AI) devices*. Type 2 or Type 3 devices embedded in mobile physical objects are often called robots.[43] As we explain below, data produced by AI devices presents unique problems when used for forensic purposes.

---

33. *Id.*
34. *Id.*
35. *See* Rebecca Darin Goldberg, *You Can See My Face, Why Can't I? Facial Recognition and Brady*, COLUM. HUM. RTS. L. REV. 261, 265 (2021).
36. *See, e.g.*, Shandra Earney, *What Are the Benefits of Smart Video Doorbells for End Users?,* XAILIENT (Jan. 25, 2023), https://xailient.com/blog/what-are-the-benefits-of-smart-video-doorbells-for-end-users/; Google Nest Help, https://support.google.com/googlenest/answer/9268625?hl=en (last visited July 5, 2023).
37. In this article we will primarily be referring to devices built into automobiles to assist (and supervise) the driver but for the general discussion the evidentiary potential of other Type 2 devices should be kept in mind. *See Fitness Trackers-Statistics & Facts*, STATISTA, https://www.statista.com/topics/4393/fitness-and-activity-tracker/#topicOverview (last visited Oct. 18, 2023); Mark Davis, *How Does Google Earth Work?*, LIVE SCIENCE (May 17, 2019), https://www.livescience.com/65504-google-earth.html; *Satellite Navigation - GPS - How it Works*, FED. AVIATION ADMIN. (last updated June 24, 2022), https://www.faa.gov/about/office_org/headquarters_offices/ato/service_units/techops/navservices/gnss/gps/howitworks.
38. *See, e.g., Roomba j9 + Robot Vacuum*, IROBOT (last visited Oct. 18, 2023), https://www.irobot.com/en_US/roomba-j9plus-robot-vacuum/J955020.html.
39. *Artificial Intelligence: What it is and Why it Matters*, SAS, https://www.sas.com/en_us/insights/analytics/what-is-artificial-intelligence.html#:~:text=AI%20adapts%20through%20progressive%20learning,product%20to%20recommend%20next%20online (last visited Oct. 18, 2023).
40. *Id. See also* Earney, *supra* note 36.
41. Corbo, *supra* note 28.
42. For details on the challenge to optimize safety-relevant and other systems based on the data collected during use, *see* WALTHER WACHENFELD & HERMANN WINNER, THE NEW ROLE OF ROAD TESTING FOR THE SAFETY VALIDATION OF AUTOMATED VEHICLES 425–30 (Daniel Watzenig & Martin Horn eds., 2017).
43. Robotnik, *The Rise of Machine Learning Robots: Explore Machine Learning in Robotics*, ROBOTNIK (June 15, 2023), https://robotnik.eu/the-rise-of-machine-learning-robots-explore-machine-learning-in-robotics/#:~:text=MACHINE%20LEARNING%20ROBOT%3A%20DEFINITION%20AND%20FUNCTIONS&text=A%20machine%20learning%20robot%20is,based%20on%20what%20it%20learns.

C. A Comparative Approach to Device Evidence

Although many people regard sophisticated devices as reliable sources of information, sometimes even more so than human beings,[44] their process of data gathering and generation is prone to errors,[45] as is the interpretation of that data.[46] Since devices operate differently from human brains and accomplish tasks differently,[47] the task of vetting device evidence raises intricate questions. New methods, benchmarks, and substantive criteria may have to be established to verify the accuracy and reliability of the operation of devices and of the data they produce.[48]

In recent years, several published studies discuss similar questions. As early as in 2007, Erin Murphy distinguished first generation (e.g., handwriting, ballistics, hair and fiber analysis) from second generation (e.g., DNA sample testing, data mining, electronic location scanning) device evidence.[49] Evaluation of second generation evidence, she wrote, requires specialized knowledge but is seemingly more scientific and, therefore, credible.[50] In a later article, Murphy argued that the safeguards inherent in the adversarial process are not well-suited to ensuring the integrity of complex technology-based forensic evidence because adequate evaluation of the accuracy and reliability of such devices needs to be done outside the courtroom rather than at trial.[51] Other scholars, building on Murphy's analysis, have demonstrated the transformation of criminal justice through the use of new technologies,[52] claiming that forensic evidence is increasingly becoming opaque for the defense, the trier of fact, and the public.[53] Consequently, traditional safeguards of the adversarial process are becoming less effective in the digital age[54] and are in need of redefinition.[55]

In 2017, Andrea Roth called for a coherent framework for conceptualizing and regulating "machine testimony," and outlined a taxonomy for such evidence along with new ideas for establishing safeguards for reliability.[56] Her proposals include testing reliability through front-end design and operation protocols, establishing new rules for pre-trial

---

44. Aleš Završnik, *Algorithmic Justice: Algorithms and Big Data in Criminal Justice Settings*, 18 EUR. J. OF CRIM. 623, 635 (2021).

45. For an overview of possible bias, "data obesity," and non-robust models, *see generally* CATHY O'NEAL, WEAPONS OF MATH DESTRUCTION (2016); Mireille Hildebrandt, SMART TECHNOLOGIES AND THE END(S) OF LAW: NOVEL ENTANGLEMENTS OF LAW AND TECHNOLOGY 34 (2015).

46. *See* Andrea Roth, *What Machines Can Teach Us about "Confrontation"*, 60 DUQ. L. REV. 210, 215 (2022*)*; Brandon L. Garrett et al., *Judges and Forensic Science Education: A National Survey*, 321 FORENSIC SCIENCE INT'L 1, 1 (2021).

47. Avery Hurt, *AI and the Human Brain: How Similar Are They?*, DISCOVER: TECH. (Jan. 14, 2023, 9:00 AM), https://www.discovermagazine.com/technology/ai-and-the-human-brain-how-similar-are-they.

48. *See* Roth, *supra* note 46, at 217–26.

49. *See* Erin Murphy, *The New Forensics: Criminal Justice, False Certainty, and the Second Generation of Scientific Evidence*, 95 CAL. L. REV. 721, 722-26 (2007).

50. *Id.* at 7.

51. Erin Murphy, *The Mismatch Between Twenty-First-Century Forensic Evidence and Our Antiquated Criminal Justice System*, 87 CAL. L. REV. 633, 659 (2013).

52. *See* ANDREW GUTHRIE FERGUSON, THE RISE OF BIG DATA POLICING 16 (2017).

53. *See* Wexler, *supra* note 23.

54. Keith A. Findley, *Innocents at Risk: Adversary Imbalance, Forensic Science, and the Search for Truth*, 38 SETON HALL L. REV. 893, 896 (2008); Brandon L. Garrett, *Big Data and Due Process*, 99 CORNELL L. REV. ONLINE 207, 208 (2014); Christophe Champod & Joelle Vuille, *Scientific Evidence in Europe--Admissibility, Evaluation and Equality of Arms*, 9, 48 INT'L COMMENT. ON EVIDENCE 1 (2011); Sonia K. Katyal, *The Paradox of Source Code Secrecy,* 104 CORNELL L. REV. 1183, 1187 (2018); Hilary Oran, *Does Brady Have Byte? Adapting Constitutional Disclosure for the Digital Age,* 50 COLUM. J.L. & SOC. PROBS. 97, 134 (2016).

55. *See e.g.*, Roth, *supra* note 46, at 211. With regard to the current practice of plea bargaining, *see, e.g.*, William Ortman, *Confrontation in the Age of Plea Bargaining*, 121 COLUM. L. REV. 451, 482 (2021).

56. *See generally* Roth, *supra* note 8.

disclosure and access, authentication, reliability, and corroboration as well as formulating specific jury instructions.[57] Recently, Paul Grimm and his colleagues have pointed out a number of problems concerning the validity and reliability of device evidence and emphasized the importance of safeguarding its accuracy.[58]

In this article, we add to the debate by introducing a comparative element. We contrast relevant parts of American evidence law with the procedural system of Germany. Expanding upon an earlier comparative study by Sabine Gless on the advantages and drawbacks of different procedural models for vetting "robot testimony,"[59] we explain the challenges that the introduction of device evidence poses in the American and the German procedural systems and propose possible solutions inspired by the German "inquisitorial" system.

As the German criminal process relies less on party initiative in presenting trial evidence and permits, to some extent, the introduction of evidence gathered in the course of the pretrial investigation,[60] it offers new perspectives and possible lessons for the American debate. For example, the German system's option to test a device using a court-appointed expert before trial seems particularly useful when device-generated data is offered as evidence. Another aspect that might prove helpful in tackling the challenges of vetting the accuracy and reliability of device evidence is a defendant's right to request a court-appointed expert whose expertise can benefit all sides.[61] This could improve the equality of arms between prosecution and defense in an area that is heavily reliant upon expert evidence. We also suggest that the problem of vetting intractable black box types of device evidence might be resolved by developing software that is able to "enter" the processes of AI-driven devices to test their accuracy and reliability.

II. DEVICE EVIDENCE IN THE UNITED STATES

A. The Evidentiary System in the United States

In the United States, non-military federal courts are governed by the Federal Rules of Evidence.[62] Each state has its own evidentiary rules but most, with the exception of California, are based to a large extent on the Federal Rules of Evidence. [63] Given this structure, this article will use the Federal Rules of Evidence and the Federal Constitution for its sources of American evidence law.[64] Although federal and state constitutions are not generally thought to contain evidentiary provisions, they often do, including aspects

57. *Id.* at 2023–36, 2038–40.
58. Paul W. Grimm et al., *Artificial Intelligence as Evidence*, 19 NW. J. TECH. & INTELL. PROP. 9, 41–84 (2021).
59. *See generally* Sabine Gless, *AI in the Courtroom: A Comparative Analysis of Machine Evidence in Criminal Trials*, 51 GEO. J. INT'L L. 195 (2020).
60. *See* Strafprozessordnung [Code of Criminal Procedure] §§ 244, 250–56 [hereinafter STPO].
61. STPO § 244 (4).
62. *See generally,* Fredric I. Lederer, *The Military Rules of Evidence, Origin and Judicial Interpretation,* 130 MIL. L. REV. 5 (1990) (The United States Armed Forces are governed by the Military Rules of Evidence, most of which are based on the Federal Rules of Evidence. However, the Military Rules not only have codified evidentiary privileges, they are also the only rules in the United States that codify the law of interrogations, search and seizure, and eyewitness identification).
63. *See generally* CAL. EVID. CODE; LEGAL INFO. INST., *Evidence – State Laws*, https://www.law.cornell.edu/wex/table_evidence (last visited Oct. 16, 2023). There are also other jurisdictions such as the tribal courts dealing with Native-American matters. *See*, *e.g.*, *Cheyenne-Arapaho Tribes of Oklahoma [Law and Order Code]*, National Indian Law Library, https://narf.org/nill/codes/cheyaracode/evidence.html (last visited Dec. 22, 2023).
64. *See generally* FED. R. EVID.; U.S. Const.

of the Fifth Amendment and, more relevant to the present discussion, the Fourth Amendment, which may affect matters such as the ability to obtain critical data.[65] The Confrontation and Compulsory Process Clauses of the Sixth Amendment also need to be considered.[66]

Despite the tendency of American media to portray both civil and criminal cases as jury trials, many trials in the United States are non-jury "bench trials"[67] where the judges act as the fact-finder. With few exceptions, the evidence rules apply equally to jury and bench trials, including the normative exclusion of hearsay (out-of-court statements offered for the truth of the matter asserted).[68] In jury trials, judges provide jurors with instructions as to the law, but jurors need not explain their factual determinations.[69] Although trial judges can certainly issue opinions, mostly dealing with applicable law, American judges need not provide any justification for their factual determinations.

B. Admissibility of Evidence—General Requirements

When considering the admissibility of a given piece of evidence, ordinarily the following evidentiary concerns must be taken into account: Logical relevance, authentication, legal relevance, hearsay, the "best evidence" (original document) rule, and the expert testimony and scientific evidence rules.[70]

We will discuss the potential evidentiary aspects of each of the three types of devices mentioned in the Introduction. The three devices are devices that record data (Type 1), devices that can draw conclusions and act upon them based on complex IT-techniques (Type 2), and AI devices which can also vary their computer code based on their interaction with new data (Type 3).

C. Devices that Store Data (Type 1)

The simplest data devices store data but neither act on that data nor analyze or classify it.[71] Normally they are coded and rely on a rule-based system.[72]

---

65. *See generally* U.S. Const. amend. V; U.S. Const. amend. IV.

66. *See generally* U.S. Const. amend VI.

67. Jeffrey Q. Smith & Grant R. MacQueen, *Going, Going, But Not Quite Gone: Trials Continue to Decline in Federal and State Courts. Does it Matter?* 101 JUDICATURE 26, 29, 37 n.37 (2017) (In 1962, in federal courts "there were more bench trials (3,037) than jury trials (2,765). This pattern continued until 1987 when, for the first time, jury trials exceeded bench trials. Today, civil jury trials occur twice as frequently as bench trials, which have constituted less than 1 percent of total civil dispositions every year since 1998."); *see New ABA Study Explains Why Jury Trials Are Disappearing*, AM. BAR ASS'N (Dec. 28, 2020), https://www.americanbar.org/news/abanews/aba-news-archives/2020/12/report-jury-trials/ (It should be noted that the total number of trials of all types have been declining in the United States for some time).

68. *See* FED. R. EVID. 1101(a)–(b).

69. *See Juror Selection Process*, UNITED STATES COURTS, https://www.uscourts.gov/services-forms/jury-service/juror-selection-process (last visited, Oct. 26, 2023) (stating that judges instruct juries about the applicable law during a trial).

70. *See generally* FED. R. EVID. 401, 403, 407, 702–03, 801-07, 901–02, 1002. The term "legal relevance" denotes Rules that constrain admissibility because of limited probative value and/or public policy (e.g., Federal Rule of Evidence 407, Subsequent Remedial Measures). The legal relevance rule of greatest potential application is Rule 403, Excluding Relevant Evidence for Prejudice, Confusion, Waste of Time, or Other Reasons, which provides that "The court may exclude relevant evidence if its probative value is substantially outweighed by a danger of one or more of the following: unfair prejudice, confusing the issues, misleading the jury, undue delay, wasting time, or needlessly presenting cumulative evidence." FED. R. EVID. 403.

71. *See e.g.* Kevin Bonsor & Nathan Chandler, *How Black Boxes Work*, https://science.howstuffworks.com/transport/flight/modern/blackbox.htm#:~:text=Older%20black%20boxes%20used%20magnetic,came%20along%20in%20the%201990s (last visited Oct. 26, 2023).

72. Mario Grunitz, *Rule-based AI vs machine learning what's the difference?*, WE ARE BRAIN (Sept. 13, 2021), https://wearebrain.com/blog/rule-based-ai-vs-machine-learning-whats-the-difference/.

i. Relevance and Authentication

The first evidentiary requirement is relevance.[73] Under the Federal Rules of Evidence, relevant evidence is admissible unless it is inadmissible under the Constitution, a statute, another rule under the Federal Rules of Evidence, or a rule prescribed by the Supreme Court.[74] Pursuant to Federal Rule of Evidence 401,

> Evidence is relevant if:
> (a) it has any tendency to make a fact more or less probable than it would be without the evidence; and
> (b) the fact is of consequence in determining the action.[75]

Assume we have a device that records a person's heart rate per minute and a party at trial wishes to enter into evidence the device's data as it relates to a given period of time. If the issue is what the individual's heartbeat was at or during a given time, the relevance of the device's data is apparent—assuming that the device functioned accurately. In this hypothetical, it should be easy to determine whether the device is accurate via expert testimony concerning its design, manufacture, and operation, which might well include evidence of experimental trials designed to test and verify accuracy.

Once relevance is established, the next likely step will be to authenticate the data, or, phrased differently, whether it is *the* data collected by the device.[76] In a way, authentication is a form of establishing relevance.[77] If a physical item or data is not what it is alleged to be, it is irrelevant.[78] Notwithstanding this, American court practice treats authentication as a special requirement of its own.[79] Federal Rule of Evidence 901(a) declares:

> To satisfy the requirement of authenticating or identifying an item of evidence, the proponent must produce evidence sufficient to support a finding that the item is what the proponent claims it is.[80]

Accordingly, the proponent of the evidence would have to show that the data offered into evidence was the data collected by and stored in the device and that it reflects the heartbeats of the given individual.[81] But does Rule 901(a) require something more—some degree of proof of accuracy? This does not appear to be the case although in the past some courts have suggested such a requirement when dealing with new technological evidence such as audio-recorded wiretaps.[82] In a recent article, Paul Grimm, Maura Grossman and Gordon Cormack seem to suggest the same for AI-based evidence.[83]

---

73. *See* FED. R. EVID. 401.
74. FED. R. EVID. 402.
75. FED. R. EVID. 401.
76. *See* FED. R. EVID. 901.
77. FED. R. EVID. 901(a) Advisory Committee's note to the 1972 proposed rules.
78. *Id.*
79. *See* FED. R. EVID. 901.
80. FED. R. EVID. 901(a).
81. *See id.*
82. *See* Edward J. Imwinkelried, *Whether the Federal Rules Of Evidence Should Be Conceived As A Perpetual Index Code: Blindness Is Worse Than Myopia*, 40 WM & MARY L. REV. 1595, 1606 (1999); *see also* EDWARD J. IMWINKELRIED, EVIDENTIARY FOUNDATIONS 79–84 (4th ed. 1998) (discussing caller ID being used as evidence and what must be offered as a foundation for the caller ID evidence).
83. Grimm et al., *supra* note 58, at 94–95.

Concededly, the issues of authentication and reliability may sometimes merge. Assume that counsel must prove that a received email is the same as the email originally composed and transmitted, and the email author is unavailable to authenticate the received email. Federal Rule of Evidence 901(b)(9) declares that authentication may be accomplished via "evidence describing a process or system and showing that it produces an accurate result."[84] This is not to prove reliability, however, but rather to help prove that something is what it purports to be. Specifically, that here, the final email is likely to be the same as the initial transmitted one.[85]

The data to be offered in evidence would most likely be in the form of raw data, requiring an expert to interpret it.[86] The device might report "heartbeats" as a text conclusion rather than showing a number reflecting, for example, electrical signals or other means of determining the heartbeat, but the device is following basic programming by which the data from the heart must be reported as a "heartbeat." In any case, we would need an expert to explain how the device works and why the factfinder should accept the monitor's conclusion as to the number of heartbeats.[87] Other evidence rules could come into play, including the hearsay[88] and "best evidence"[89] rules. Notwithstanding this, if the evidentiary requirements above have been met, it will usually suffice to permit the admission into evidence of basic data obtained from an electronic device.[90] However, given that the American legal system is based on jury trials, there is often a concern that lay jurors will over or under value evidence and that as a result the judge should be able to exclude such evidence.[91] The "legal relevance" rules, as they are sometimes referred to by academics, limit or prohibit admission of evidence because of concerns about its probative value and/or reasons of public policy.[92] The primary such rule is Rule 403, which states:

> The court may exclude relevant evidence if its probative value is substantially outweighed by a danger of one or more of the following: unfair prejudice, confusing the issues, misleading the jury, undue delay, wasting time, or needlessly presenting cumulative evidence.[93]

The Rule subsumes the common law "unfairly prejudicial" objection to evidence.[94] Under that rule, defense counsel in a criminal homicide case could, for example, object to close-up photos of the body of the deceased victim showing 53 brutal stab wounds.[95] The key to Rule 403 is its restriction on evidence when its probative value is *substantially* outweighed

---

84. FED. R. EVID. 901(b)(9).
85. *See, e.g.* KENNETH S. BROUN ET AL., MCCORMICK ON EVIDENCE 398 (6th ed. 2006) (stating that "even perceived errors in the output are said to go to the weight of the evidence, not its admissibility.").
86. *See* FED. R. EVID. 702(a).
87. *Id.*
88. FED. R. EVID. 801, 802 (dealing with out-of-court statements offered for their truth). Normatively, such statements are inadmissible. However numerous exceptions exist to the rule. *See* FED. R. EVID. 801(d), 803, 804, 807.
89. This rule, despite the name, is actually limited to creating a requirement for producing originals when proving the contents of documents or their equivalents. *See e.g.*, FED. R. EVID. 1001, 1002. Interestingly, Federal Rule of Evidence 1001(d) declares that "[f]or electronically stored information, 'original' means any printout — or other output readable by sight — if it accurately reflects the information."
90. *See* FED. R. EVID. 1001(d).
91. *See* FED. R. EVID. 403 Advisory Committee's note to the 1972 proposed rules.
92. FED. R. EVID. 401, 403.
93. FED. R. EVID. 403.
94. FED. R. EVID. 403 advisory committee's note to the 1972 proposed rules.
95. *See, e.g.*, United States v. Matthews, 13 M.J. 501, 517–18 (A.C.M.R. 1982), reversed in part on other grounds, 16 M.J. 354 (C.M.A. 1983). Note that in *Matthews*, the Court held that the pictures were properly admitted. *Matthews*, 13 M.J. at 518.

by factors such as unfair prejudice.[96] At least in the United States, judges and jurors may give technology-derived evidence undue weight, due to "automation bias."[97] Accordingly, otherwise admissible evidence could be inadmissible if such bias is thought to substantially outweigh the probative value of the evidence.[98] These rules also apply to bench trials without juries.[99]

ii. Expert Testimony

In the United States, most judges and jurors lack the degree of knowledge and expertise necessary to understand scientific, technological, or medical evidence, to address only a few specialized subjects.[100] Accordingly, the United States legal system permits the use of subject matter experts when their testimony can "help" the trier of fact, as set forth in Federal Rule of Evidence 702.[101] Admissibility of device evidence customarily will require an explanation of how the device works and whether data obtained from it is accurate and reliable.[102] That explanation ordinarily would be furnished by expert witnesses.[103]

The common law system is based on in-court presentation of evidence and emphasizes the perceived utility of cross-examination.[104] Accordingly, device evidence in the United States should be subject to a thorough open-court inquiry. Sophisticated device data may be very difficult to understand and verify.[105] In light of the nature of that evidence, and the Supreme Court's requirements in *Daubert v. Merrell Dow Pharmaceuticals,*[106] device evidence ought to be tested thoroughly by expert testimony. The American adversary system compels the use of experts who are called to substantiate their party's

96. FED. R. EVID. 403.
97. *See, e.g.,* Danielle Keats Citron, *Technological Due Process*, 85 WASH. U. L. REV. 1249, 1271–72 (2008); *see also* Kathleen L. Mosier et al., *Automation Bias: Decision Making and Performance in High-Tech Cockpits*, 8 INT'L J. AVIATION PSYCH. 47, 59 (1998) (discussing the effects of automation bias for pilots and how it affects their decision making while operating an aircraft depending on the level of self-accountability that they feel while flying); Mary Cummings, *Automation Bias in Intelligent Time Critical Decision Support Systems*, AM. INST. AERONAUTICS & ASTRONAUTICS 1ST INTELLIGENT SYS. TECH. CONF. 20–22 September 2004, 6313 (discussing the effects of automation bias in aviation with respect to computer assisted route planning, event diagnosis and action, and time sensitive resource allocation); Murphy, *supra* note 49, at 737, 757 (discussing prosecutors' preference for evidence where "proof of scientific certainty is readily available."); Patrick W. Nutter, *Machine Learning Evidence: Admissibility and Weigh*t, 21 U. PA. J. CONST. L. 919, 949 (2019) (suggesting that "the manner in which the Sixth Amendment requires expert witnesses to testify on drug analysis evidence may provide a framework for how machine learning experts would be required to testify in-person and be subject to cross-examination.").
98. *See* FED. R. EVID. 403.
99. *See* FED. R. EVID. 1101(a)–(b).
100. *See* FED. R. EVID. 702 advisory committee's note to the 1972 proposed rules.
101. FED. R. EVID. 702(a).
102. *See* FED. R. EVID. 702.
103. *See* FED. R. EVID. 702(a).
104. *See, e.g.* Crawford v. Washington, 541 U.S. 36, 43 (2004).
105. Vanessa Buhrmester et al., *Analysis of Explainers of Black Box Deep Neural Networks for Computer Vision: A Survey*, 3 *MACH. LEARNING AND KNOWLEDGE EXTRACTION* 966, 984 (2021); Cynthia Rudin, *Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead,* 1 NATURE MACH. INTEL. 206, 207–09 (2019).
106. 509 U.S. 579, 593–95 (1993); *see* FED. R. EVID. 702 Advisory Committee notes on the 2000 amendments.

perspective.[107] The potential use of device evidence will therefore require substantial expert analysis before trial in addition to their trial testimony.[108]

> Pursuant to Federal Rule of Evidence 702 (Testimony by Expert Witness),
> A witness who is qualified as an expert by knowledge, skill, experience, training, or education may testify in the form of an opinion or otherwise if:
> (a) the expert's scientific, technical, or other specialized knowledge will help the trier of fact to understand the evidence or to determine a fact in issue;
> (b) the testimony is based on sufficient facts or data;
> (c) the testimony is the product of reliable principles and methods; and
> (d) the expert has reliably applied the principles and methods to the facts of the case.[109]

Federal Rule of Evidence 702 is based upon the U.S. Supreme Court's decision in *Daubert* and is intended to ensure the validity of the expert's testimony.[110] *Daubert* implicitly requires that the science, medicine, or technology underlying the expert's testimony is valid and reliable.[111] In *Daubert,* the Court focused upon the admissibility of scientific expert testimony.[112] It pointed out that such testimony is admissible only if it is both relevant and reliable[113] and held that the Federal Rules of Evidence "assign to the trial judge the task of ensuring" reliability.[114]

In the later case of *Kumho Tire Co. v. Carmichael*, the Supreme Court expanded *Daubert* to include technology-based evidence.[115] A careful application of the Supreme Court's approach in *Daubert v. Merrell Dow Pharmaceuticals* is therefore likely to suffice for admissibility of ordinary device data.[116]

But given the complexity of device evidence and its increasing importance, the American legal system is not well designed to ensure its efficient and reliable use.[117] Although Federal Rule of Evidence 706 permits a federal judge in a federal case to appoint experts, the general custom in American courts is that parties obtain their own experts.[118] This custom unavoidably leads to the introduction of partisan testimony.[119] Experts identified with a given party are likely to be less credible in the view of the factfinder, whether

---

107. *See* FED. R EVID. 702.
108. *See* Buhrmester et al., *supra* note 105, at 984 (discussing the difficulties in understanding data obtained from black box deep neural networks); *see also* FED. R. EVID. 702 (stating than an expert may testify in the form of an opinion if their specialized knowledge will allow them to help the jury understand the evidence).
109. FED. R. EVID. 702.
110. *See* FED. R. EVID. 702 Advisory Committee notes on the 2000 amendments.
111. 509 U.S. at 594–95 (1993).
112. *Id.* at 582.
113. *Id.* at 592–93.
114. *Id.* at 597.
115. *Kumho Tire Co.*, 526 U.S. 137, 147, 149-50 (1999).
116. *See id.*
117. *See* JOE S. CECIL & THOMAS E. WILLGING, COURT-APPOINTED EXPERTS: DEFINING THE ROLE OF EXPERTS APPOINTED UNDER FEDERAL RULE OF EVIDENCE 706 8 (1993).
118. *Id.*
119. Adam Liptak, *In U.S., Expert Witnesses Are Partisan*, N.Y. TIMES (Aug. 8, 2008), https://www.nytimes.com/2008/08/12/us/12experts.html.

judge or jury.[120] Given that counsel will not hire experts who disagree with his or her client's theory of the case and, indeed, may search long and hard for an expert who will concur with their theory, the American battle of partisan expert witnesses hardly inspires confidence in the quality of expert testimony.[121]

Moreover, expert testimony does not come without a price tag, which leads to the problem of the reduced availability of qualified experts for indigent parties, especially defendants in criminal cases.[122] With respect to payment of experts, Rule 706(c) provides that experts are "entitled to a reasonable compensation, as set by the court."[123] In criminal cases, the compensation is payable from any funds that are provided by law, and a defendant unable to afford experts may apply to the court for assistance;[124] but the success of such an application in any given case is doubtful.

D. Devices that Evaluate Data (Type 2)

We can now advance to devices that not only collect and store data but also draw conclusions from that data. Determining the accuracy of "evaluative data"[125] from devices can be especially difficult as it is the result of a device's autonomous assessment of its environment. The genesis of such data cannot be completely understood by humans due to the complexity of the algorithms and/or the impact of machine learning. For example, not all information fed into a drowsiness alert system (e.g., lane marking, road condition, lighting conditions) is stored.[126] Secondly, due to the use of machine learning, humans cannot understand how and why a device acted on given data.[127] This notorious black box problem cannot even be solved by introducing an expert to engage with the device, unless costly and sophisticated methods like reverse data engineering can be used.[128]

To see how his problem could present itself, let's return to the "testimony" of the BMW automobile that began this article. As the reader will recall, the automobile concluded that its driver was responsible for the collision that followed the automobile's warning that it was getting close to the automobile in front of it. Imagine how part of the cross-examination might go—if the automobile used Natural Language Processing to "understand" and respond to the human counsel's questions.

i. Questioning the BMW—Part 2

Q: How did you know how close you were to the automobile in front of you and how fast you were approaching it?

---

120. *Id.*; *see* CECIL & WILLGING, *supra* note 119, at 13, 27, 50.
121. Liptak, *supra* note 121.
122. *See* CECIL & WILLGING, *supra* note 119, at 5.
123. FED. R. EVID. 706(c).
124. Ake v. Oklahoma, 470 U.S. 68, 74 (1985) (holding that a murder defendant had a constitutional right of access to a competent psychiatrist when his sanity was in question).
125. *See* Emily Silverman et al., *Robot Testimony? A Taxonomy and Standardized Approach to Evaluative Data in Criminal Proceedings* (Sabine Gless & Helena Whalen-Bridge eds., forthcoming 2024).
126. Sabine Gless et al., *Ca(r)veat Emptor: Crowdsourcing Data to Challenge the Testimony of In-Car Technology*, 62 JURIMETRICS 285, 289, 294 (2022).
127. *See* Gless, *supra* note 59, at 211.
128. *Id.*

A. I am equipped with both forward-facing radar and a video camera. My computer chips are programmed to accurately measure distance from these devices and for me to alert the driver if it appears that the driver doesn't recognize a likely collision risk.

Q: When were these systems last checked for accuracy?

A: I do not know; they should have been checked during my last major maintenance.

Q: Are you able to explain the algorithm that determines collision risk?

A: No—but the original algorithm is available from BMW.

If our BMW is measuring and reacting to exterior conditions such as weather, road surface and the like, it is dependent on its sensors.[129] But were the sensors accurate? It is probable that so long as the sensors appear operational, the BMW relies on dealer maintenance of the sensors and, as our BMW cross-examination notes, the system may not recognize what was done in maintenance, why, or to what effect. Determining why a device interpreted data the way it did, drew conclusions from that data, and then determined how best to act on that data can thus be difficult if not impossible.

Let us return to the simple example of a heart monitoring device. A more sophisticated device, including current Apple watches or advanced Fitbits, might take the equivalent of an electrocardiogram ECG and warn of heart conditions such as arrhythmia.[130] If a person were to testify in court that she had arrhythmia based on what her watch reported, the accuracy of that information would require testimony from one or more experts.[131] The proponent of the evidence would have to establish:

> - the soundness of the underlying science and technology used in the watch;
> - the soundness of the design of the watch, including both hardware and software; and
> - the accuracy and reliability of the actual hardware and software implementation, including its results.

Symptoms often are not certain proof of a given bodily condition.[132] Accordingly, counsel would have to present expert testimony as to how the programming treats that uncertainty of the device's assessment—most likely by embodying a probability design.[133] At this point, the accuracy and reliability of the algorithm comes into play. The algorithm is written by fallible human beings, who may also have improperly classified the data relied upon by the algorithm.[134] Moreover, data selected for training or programming devices

---

129. *See* Gless et al., *supra* note 128, at 286, 288, 289.
130. Apple Support, *Take an ECG with the ECG App on Apple Watch*, APPLE INC., https://support.apple.com/en-us/HT208955 (last visited Nov. 6, 2023).
131. *See* Gless, *supra* note 59, at 211–12.
132. Marianne Rosendal et al., *"Medically Unexplained" Symptoms and Symptom Disorders in Primary Care: Prognosis-Based Recognition and Classification*, BMC FAMILY PRACTICE, Feb. 7, 2017, at 2.
133. Fredric I. Lederer, *Problematic AI – When Should We Use it?,* HARV. ADVANCED LEADERSHIP INITIATIVE SOC. IMPACT REV. (2022); Buhrmester et al., *supra* note 105, at 969.
134. *See generally,* Citron, *supra* note 97.

may be biased or delusive.[135] For instance, if a drowsiness detection system is trained solely on data generated during test drives with athletic Caucasian males, it has a "white guy problem"[136] and might conclude that a female of Asian descent is "drowsy" simply because of her sitting position and eye shape.[137]

The capability of Type 2 devices to autonomously draw conclusions from data and to act on them comes at the price of a black box problem.[138] As a machine, the device can neither critically reflect on its assessments nor provide information on possible misunderstandings.[139] And human beings, including IT experts testifying at a trial, may not be able to determine how a device made a given decision or functioned the way it did,[140] especially if the device had been trained with machine learning techniques of such complexity that they are beyond human understanding.[141] Sometimes devices are afforded multiple ways of interpreting data and can "choose" what seems to them the most accurate way to interpret and label the data.[142]

Therefore, while cross-examination in the adversarial trial and rights of discovery and confrontation in the inquisitorial trial have been crucial for vetting the credibility of human witnesses, they seem ineffective when applied to Type 2 devices.[143] It is thus difficult to establish the accuracy and reliability of device-generated evidence.[144] In the case of rule-based systems, experts can explain how the system collects and processes data and comes to a result.[145] This is much more difficult in complex systems that process a plethora of data and rely on a training data set not known to the public.[146] An expert cannot fully trace the device's path from the collection of information to an evaluative assessment.[147] If, for example, a combined lane-keeping assistant and drowsiness detection system relies on variable reference points that defy standardized measurement (such as the driver's body tension and the movements of the driver's eyelids) and then evaluates them independently, even experts may be unable to verify the correct working of the system.[148]

E. Devices with the Capacity for Self-Modification (Type 3)

Devices that have undergone specific, highly complex machine-learning techniques and can modify their own operations based on their "experience" pose the greatest challenge.[149] The option of self-optimization that involves an adaption of the code

---

135. Vivek Khetan, *Bias in Machine Learning Algorithms,* TOWARDS DATA SCIENCE (Apr. 6, 2019), https://towardsdatascience.com/bias-in-machine-learning-algorithms-f36ddc2514c0.
136. Sabine Gless, Xuan Sharon Di & Emily Silverman, *Ca(r)veat Emptor: Crowdsourcing Data to Challenge the Testimony of In-Car Technology*, 62 JURIMETRICS 285, 291 (2022).
137. *See generally* Kristin N. Johnson, *Automating the Risk of Bias,* 87 GEO. WASH. L. REV. 1214 (2019); Cathy O'Neil, Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy (2016).
138. *See* Gless, *supra* note 59, at 211.
139. Khetan, *supra* note 137.
140. Buhrmester et al., *supra* note 105, at 966, 984.
141. Rudin, *supra* note 105, 206–07.
142. *Id.* at 225.
143. Roth, *supra* note 46, at 211.
144. *Id.*
145. *See* Gless, *supra* note 59, at 211–12.
146. *See* Gless et al., *supra* note 128, at 294.
147. *Id.*
148. *Id.*
149. Brenden M. Lake et al., *Building Machines That Learn and Think Like People*, 40 BEHAV. & BRAIN SCI. 16–17 (2016).

responsible for generating data obviously makes it even more difficult for a court to determine why the device did what it did, and whether the evidence presented is reliable. Let us return to our—now futuristic—BMW:

i. BMW Testimony—Part 3

Q: You say, "original" algorithm; isn't that what you used at the time of this collision?

A: No, my algorithm can optimize safety features automatically as I operate so as to make my monitoring tools work as accurately as possible.

Q: Can you tell us when it changed and for each change how and why it changed?

A: No; I was not designed to be able to do that.

Q: Let's go on to another matter, do you know why your driver allegedly went faster and tried to pass the car before you?

A. No.

Q: So, if your driver actually did that, there could have been a legitimate, perhaps impelling safety reason to do so?

A: I do not know; I am not programmed to consider all, especially extraordinary, human actions.

This final part of the interrogation of the talking BMW opens up a new vista: The car's advanced form of machine learning is characterized by algorithms that are able to self-optimize and modify their programming during operation. If a system is designed to maximize human productivity in a factory, for example, it might correlate productivity data with ambient temperature in various parts of the factory and modify that temperature until workers reach maximum productivity—a result that might vary by day, season, or time of day.

Due to their specific learning model and their heuristics, Type 3 AI devices face several typical limitations and sources of error. False results may occur due to inadequate programming, software design or training, mistaken self-learning,[150] or external interference with the acquisition or processing of data.[151] AI devices can also deliver inappropriate responses if they proceed on the basis of incorrect premises or misinterpret the data, for example, if a system was exposed to large amounts of possibly incorrect and biased data, such as that found throughout the Internet.[152]

---

150. *See* Buhrmester et al., *supra* note 105, at 967 (stating that incorrect training data can lead to false results, if for instance a dog or wolf classifier is trained on pictures when most of the photos for the training set of wolves are taken on days with snowy weather, while the dog images are taken on sunny days, the classifier will turn out to be just a good snow detector).

151. *See* Cao et al., *Adversarial Sensor Attack on LiDAR-Based Perception in Autonomous Driving*, *in* PROCEEDINGS OF THE 2019 ACM SIGSAC CONF. ON COMPUT. AND COMMC'N SEC., 2267, 2267–69 (2019).

152. Chad Boutin, *There's More to AI Bias Than Biased Data*, *NIST Report Highlights*, NIST (Mar. 16, 2022), https://www.nist.gov/news-events/news/2022/03/theres-more-ai-bias-biased-data-nist-report-highlights.

But it may be difficult if not impossible to discover such defects. Absent a full record of every aspect of an AI system's training, [153] its programming based partly on a "self-teaching" machine learning process, and the information processed in its ongoing interface with the world, it is currently impossible to establish why an AI device reacts in a given way to a specific situation.[154] In the case, for example, of a lane-keeping assistant that processes a plethora of environmental information (line marking, road texture, light and shadows) with every mile driven and in every situation encountered, it would be necessary to retroactively determine all information gathered and processed by the algorithms in place to determine whether an accident was due to a fault of the driver or to the car's misguided (self) programming.[155] But, due to the magnitude of the data involved, devices normally do not store the information gathered by sensors, which means that crucial data is missing when a court has to decide on someone's civil or criminal liability.[156] The "testimony" (in whichever way introduced at a criminal trial) of an AI device is not, therefore, reliable evidence for answering the critical factual questions of a case. Such evidence would thus at present be irrelevant or if relevant subject to an inadmissibility determination under the legal relevance rule, Federal Rule of Evidence 403.[157]

Before we consider how Germany deals with the challenges of device evidence, we will briefly review the possible impact of some relevant provisions of American constitutional law on proffering device evidence as proof.

F. Constitutional Constraints

Under the Federal Constitution, in criminal cases the rights to confrontation, compulsory process, due process, and equal protection are facially relevant.[158] Emerging changes triggered by new technology have led to lively debates, for instance on how the Bill of Rights ought to be interpreted in light of modern technology.[159]

i. Fourth Amendment: Privacy

An applicable right to privacy can limit the type and amount of data extracted from a device.[160] This can be of importance in circumstances such as the seizure of the contents of one's smart phone via data captured by an automobile's entertainment system.[161]

---

153. *See* Andres J. Ramirez & Betty HC Cheng, *Design Patterns for Developing Dynamically Adaptive Systems*, PROC. OF THE 2010 ICSE WORKSHOP ON SOFTWARE ENG'G FOR ADAPTIVE AND SELF-MANAGING SYS. 49, 52–54 (2010).
154. *See* Gless, *supra* note 59, at 211–13; Fredric I. Lederer, *Problematic AI – When Should We Use it?,* HARV. ADVANCED LEADERSHIP INITIATIVE SOC. IMPACT REV. (2022).
155. *See* Gless et al., *supra* note 128, at 289.
156. *Id.* at 285, 289.
157. FED. R. EVID. 403.
158. Garrett, *supra* note 54, at 207, 212.
159. Findley, *supra* note 54, at 944, 948, 951; Garrett, *supra* note 54, at 207, 208; Murphy, *supra* note 51, at 635–39; Oran, *supra* note 54, at 98–99; Wexler, *supra* note 23, at 1352–53; Ortman, *supra* note 55, at 454–55; Roth, *supra* note 46, at 210–11.
160. *See* STEPHEN J. SCHULHOFER, MORE ESSENTIAL THAN EVER: THE FOURTH AMENDMENT IN THE TWENTY-FIRST CENTURY 115–43 (Geoffrey R. Stone et al. eds., 2012).
161. *See* Gless et al., *supra* note 128, at 293.

Subject to applicable exceptions, the Fourth Amendment prohibits the government from searching and/or seizing without a judicial warrant[162] when a person has a reasonable expectation of privacy in the location to be searched or the item to be seized.[163] This expectation of privacy can extend to devices containing data, and some modern devices can indeed be a treasure trove of highly private data. Recognizing the particular vulnerability of the individual with regard to such data carriers, the U.S. Supreme Court in 2014 held that an exception to the warrant requirement that permits searches of items incident to a lawful arrest does not apply to smart phones because of the amount of personal data typically stored in them.[164] Distinguishing past decisions involving searches of the person incident to arrest, the Supreme Court noted that an arrested person's loss of privacy following arrest does not affect cell phone data, and that cell phones "place vast quantities of personal information literally in the hands of individuals."[165] Notwithstanding the Court's recognition of the importance of data to the modern individual, its other cases dealing with the "third party doctrine" suggest substantial limits on Fourth Amendment protection of data under present conditions of data storage. Pursuant to the third-party doctrine, a person cannot claim a reasonable expectation of privacy in data that has been transmitted or made accessible to others.[166] In its 2018 decision in *Carpenter v. U.S.*[167], the U.S. Supreme Court recognized the risk to privacy posed by application of the third party doctrine to electronic data but failed to resolve the general problem. The Court held that a Fourth Amendment "search" occurs when a state agency requests historical cell site location information ("CSLI") concerning a private smartphone from a commercial wireless carrier.[168] However, the normal exception to the warrant requirement for time sensitive exigencies continues.[169] Overall, *Carpenter* suggests that the Supreme Court clearly recognizes that data is different and that data searches and seizures will require reevaluation of traditional precedents.[170]

Accordingly, it is possible that the Fourth Amendment may protect, to some degree, the holder of data against a government search or seizure of that data, subject to numerous exceptions.[171]

---

162. It should be noted, however, that even if a data search or seizure requires a warrant, warrants are not difficult to obtain. An empirical study found that most magistrates routinely approved warrant requests. RICHARD VAN DUIZEND ET AL., THE SEARCH WARRANT PROCESS: PRECONCEPTIONS, PERCEPTIONS, AND PRACTICES 35 (Carolyn McMurran ed., 1985). The study concluded that "[t]he average length of the magisterial review . . . was two minutes and forty-eight seconds. The median time was two minutes and twelve seconds." *Id*. at 26. Some magistrates even authorized searches, knowing them to be unlawful. FREDRIC I. LEDERER, FUNDAMENTAL CRIMINAL PROCEDURE 84 (2022).

163. Katz v United States*,* 389 U.S. 347, 360 (1967).

164. Riley v. California, 573 U.S. 373, 403 (2014).

165. *Id.* at 386.

166. Erin Murphy, *The Case against the Case for Third-Party Doctrine: A Response to Epstein and Kerr*, 24 BERKELEY TECH. L.J. 1239, 1239 (2009).

167. 138 S. Ct. 2206 (2018).

168. Carpenter v. United States, 138 S. Ct. 2206, 2217 (2018). For an earlier discussion of the issue, see the concurring opinion of Justice Sotomayor in *United States v. Jones,* 565 U.S. 400, 413–14 (2012).

169. Carpenter, 138 S. Ct. at 2222–2223.

170. *Id.* at 2217.

171. *See* Katherine J. Strandburg, *Home, Home on the Web and Other Fourth Amendment Implications of Technosocial Change*, 70 MD. L. REV. 614, 622, 650–79 (2011).

ii. Fifth Amendment: Due Process

In criminal cases, both constitutional law and rules such as Rule 16 of the Federal Rules of Criminal Procedure require that the prosecution supply the defense with significant information, including anything known to the prosecution that might be beneficial to the defense case. [172] Accordingly, the defense should have advance notice of any device evidence that may be of use for the defense.[173] Discovery rules require providing criminal defendants with information about the design of a device, its sensors, and its basic programming.[174] With regard to "interpretable" or "explainable" data,[175] however, the black box effect of certain machine learning schemes prevents defendants from obtaining a meaningful explanation of causation.[176] Moreover, granting defense lawyers access to device-generated information in the hands of the prosecution during discovery is a poor substitute for the ability to examine device evidence and the device directly, and its value is dependent upon the prosecution's ability to recognize the potential defense value of the information in its possession.[177]

Even where the U.S. adversarial system provides the defense with the right to discover prosecution evidence, defense lawyers will be at a loss when they attempt to test devices and to interpret their operation.[178] And, as Shakespeare had Hamlet exclaim, "Ay, there's the rub."[179] To determine the validity of device data, defense lawyers need to have access to competent experts, but experts in this field are rare and expensive.[180] In important civil cases, a poor plaintiff may be able to obtain financial support from litigation support firms that invest in civil cases for a share in the proceeds, thus permitting the hiring of otherwise unaffordable experts, but this is unavailable in criminal cases.[181]

iii. Sixth Amendment

a. Right to Confrontation

The Sixth Amendment to the United States Constitution provides that "*In all criminal prosecutions, the accused shall enjoy the right . . . to be confronted with the witnesses against him".*[182] According to the United States Supreme Court in *Crawford v. Washington*,[183] the confrontation right was intended to reject the ancient European inquisitorial system of relying on pretrial written evidence and to require in-court testimony.

---

172. *See, e.g.,* United States v. Augurs, 427 U.S. 97, 112–14 (1976) (absent a specific defense request, the prosecution must disclose only that evidence which creates a reasonable doubt that would not otherwise exist); Brady v. Maryland, 373 U.S. 83, 87–88 (1963); FED. R. CRIM. P. 16(a). *See also* Kyles v. Whitley, 514 U.S. 419, 435 (1995) (the general test for non-disclosure with or without a defense request is whether there would be a reasonable probability of an acquittal had the information been disclosed by the prosecution).

173. FED. R. CRIM. P. 16(a).

174. FED. R. CRIM. P. 16(a)(1)(E) (requiring the prosecution to permit the defense to inspect and copy "documents, data, photographs, [and] tangible objects").

175. *Cf.* Buhrmester et al., *supra* note 105, at 972.

176. Roth, *supra* note 8, at 1989–90.

177. *Cf.* for a general assessment of Murphy, *supra* note 51, at 647–50.

178. Roth, *supra* note 8, at 1980.

179. WILLIAM SHAKESPEARE, HAMLET act 3, sc. 1, l. 65.

180. *See, e.g.*, William A Ramsey, *Court Issues Decision Clarifying Reasonableness of Expert Witness Fees*, BARRETTMCNAGNY: APPELLATE LAW BLOG (last visited Nov. 6, 2023), https://www.barrettlaw.com/blog/appellate-law/court-issues-decision-clarifying-reasonableness-of-expert-witness-fees.

181. *See, e.g.*, Jarrett Lewis, *Third-Party Litigation Funding: A Boon or Bane to the Progress of Civil Justice?*, 33 GEO. J. LEGAL ETHICS, 687, 687–88 (2020).

182. U.S. CONST. amend. VI.

183. 541 U.S. 36, 44–45 (2003). Note that Federal Rule of Evidence 801(b) defines for hearsay purposes a declarant as "the person who made the statement." *See generally* Jeffrey Bellin, *The Incredible Shrinking Confrontation Clause*, 92 B.U. L. REV. 1865 (2012); FED. R. EVID. 801(b).

In substance, the Supreme Court has interpreted the Confrontation Clause to prohibit the use of prosecutorial "testimonial" hearsay—out-of-court statements offered for the truth of the matter asserted with the expectation that they would be used at trial.[184] Instead, witnesses are to testify at trial to what they personally observed.[185]

Although the major questions relating to the confrontation clause today deal with the scope of *Crawford* and the degree to which it limits remote prosecution testimony,[186] one can ponder the potential impact of the confrontation clause on device evidence. Assume the universe of Isaac Asimov's sentient, intelligent, and independent robots.[187] If such a robot were to be called to testify by the prosecution in an American criminal case, would the confrontation clause apply? Would such a robot be a "person" or a "device" for purposes of the Bill of Rights? It seems clear that American courts will not treat today's and tomorrow's devices as "persons." Given that the conservative Supreme Court justices are originalists, striving to determine how the Framers understood and intended constitutional provisions to be used, they must be expected to apply the Confrontation Clause only to statements made by human beings.[188] But even if an intelligent device were to be treated by law as the equivalent of a human being,[189] absent a fundamental change in machine learning technology even a human-seeming "witness" cannot meaningfully be cross-examined. Its reliability therefore cannot be established in the usual procedural way. In pre-*Crawford* cases such as *Ohio v. Roberts,*[190] the U.S. Supreme Court emphasized the need to demonstrate the accuracy of evidence that is not subject to cross-examination. Although abandoned by *Crawford,* the policy could easily be applied to device evidence. Applying that rationale to a robot's "testimony," it is apparent that reliability cannot be established, thus making the "testimony" violative of the confrontation clause, absent the existence of special tests that could serve as a functional equivalent of cross-examination.[191]

In the United States, a right to "confront" device evidence could be created by statute or perhaps in criminal cases be founded in due process concerns. But what would such a right mean? Current discovery should require providing criminal defendants with the ability to obtain information about the design of a device, its sensors, and its basic programming.[192] Subject to the development of "interpretable" or "explainable" AI,[193] however, the black box effect of machine learning prevents defendants from obtaining a meaningful explanation of causation. Even a technician who operated the

184. *Crawford*, 541 U.S. at 51–52.
185. *See generally Crawford*, 541 U.S. 36.
186. *See, e.g.,* Fredric I. Lederer, *The Evolving Technology-Augmented Courtroom Before – During – and After the Pandemic*, 23 VAND J. ENT. & TECH. L. 301, 320 n.85 (2021).
187. *See, e.g.*, Cathy Lowne & Pat Bauer, *I, Robot work by Asimov*, BRITANNICA: ARTS AND CULTURE (last updated June 16, 2023), https://www.britannica.com/topic/I-Robot.
188. *See* Bullcoming v. New Mexico, 564 U.S. 647, 651 (2011); Melendez-Diaz v. Massachusetts, 557 U.S. 305, 312 (2009).
189. Applying the 5th amendment to a robot would certainly raise problems: Could a robot take an oath or affirmation? *See* U.S. CONST. amend. V.
190. 448 U.S. 56, 57 (1980).
191. Roth, *supra* note 46, at 212.
192. FED. R. CRIM. P. 16(a).
193. *Cf.* Buhrmester et al., *supra* note 105, at 972.

device would probably be unable to provide that explanation.[194] Recognition or creation of a right to "confront" device evidence might therefore lead to the exclusion of device evidence offered by the prosecution in criminal cases.[195]

b. Right to Compulsory Process

The Sixth Amendment's Compulsory Process Clause, ("*to have compulsory process for obtaining witnesses in his favor"*) was applied in the seminal case of *Chambers v. Mississippi*[196] to grant the defendant the right to present important probative hearsay evidence even when ordinarily inadmissible under applicable state law. In our context, this could mean that the defense may call an expert on device evidence even if there is no meaningful cross-examination possible due to the black box features of the device.[197] It may even be that the device's data, including its conclusions, might be directly admissible in favor of the defense if that data were determined by the judge to be highly probative on the facts of the specific case.[198]

As a preliminary result of our brief foray into American constitutional law, we can say that the accuracy and reliability of device-generated data must be established if it is to be used for evidentiary purposes.[199] Two main approaches to reaching this goal appear feasible: A normative approach that establishes a defendant's right to have the device's accuracy and reliability checked in a meaningful way and a technological approach that relies on technical solutions for the same purpose.[200] Alternatives in that regard are certification and approval procedures for devices ("front-end design") and AI driven tools that can verify devices' findings on a case-by case basis.[201] We will address these alternatives in greater detail below. But first, let us examine the German criminal process and see what lessons it may hold for the United States.

---

194. *See generally* Brian Sites*, Rise of the Machines: Machine-Generated Data and the Confrontation Clause*, 16 COLUM. SCI. & TECH. REV. 36 (2014).

195. *See, e.g.,* U.S. CONST. amend. VI.

196. 410 U.S. 284, 284 (1973); U.S. CONST. amend. VI. In *Chambers*, a homicide case, a police officer attempting to execute an arrest warrant was attacked by a crowd. 410 U.S. at 285-86. The primary prosecution evidence of who fired at the officer was that the officer seemed to fire his riot gun at a man running down an alley, Leon Chambers, a Black man. *Id.* at 286. The officer then died from several shots in the back. *Id.* Another man confessed to shooting the officer but later retracted his confessions. *Id.* at 287-88. Chambers was tried and convicted of killing the officer. *Id.* at 285. Under state evidentiary law the defense was unable to treat the person who had confessed as an adverse witness and to cross-examine him. *Id.* at 291. That and the hearsay rule substantially hindered Chambers' defense. *Id.* at 294. The Court held that the cumulative effects of the exclusion of evidence denied Chambers due process. *Id.* at 302. '"Broadly construed, [Chambers] appears to recognize that the accused in a criminal proceeding has a constitutional right to introduce *any* exculpatory evidence, unless the state can demonstrate that it is so inherently unreliable as to leave the trier of fact no rational basis for evaluating its truth.'" EDWARD J. IMWINKELRIED, PAUL C. GIANNELLI, FRANCIS A. GILLIGAN, FREDRIC I. LEDERER & LIESA RICHTER, II COURTROOM CRIMINAL EVIDENCE 6-83 (7th ed. 2022) (citing Peter Westen, *The Compulsory Process Clause*, 73 MICH. L. REV. 71 (1974)).

197. *See Chambers*, 410 U.S. at 284.

198. *See, e.g.*, *id.* at 302.

199. Alex Nunn, *Machine-Generated Evidence*, AMERICAN BAR ASSOCIATION (July 1, 2020), https://www.americanbar.org/groups/science_technology/publications/scitech_lawyer/2020/summer/machinegenerated-evidence/.

200. *See* Gless, *supra* note 59, at 248.

201. *See* Gless, *supra* note 59, at 248–249; Roth, *supra* note 8, at 2028.

III. DEVICE EVIDENCE IN GERMANY

A. Principles of German Procedure Law Applicable to Criminal Cases

The origins of German criminal procedure lie in continental Europe's traditional inquisitorial process. In that type of process, it is the judges' task to determine the truth about a criminal incident and they are obliged to do everything within their legal authority to discover "the truth."[202] Germany still adheres to this principle today. German law thus places a great deal of responsibility on judges to determine the facts of a case. Although counsel can play an important role at the trial, it is the presiding judge who is responsible for introducing the evidence relevant to the court's findings of fact, including appointing experts.[203] A similar duty is placed upon the public prosecutor in the pre-trial phase of the process—when a complaint or other information suggests the possibility that a crime was committed, the prosecutor (and in practice, the police) investigates the matter, gathers information, and places it in a dossier.[204] If the prosecutor establishes sufficient cause for filing a formal accusation, the dossier is passed on to the trial court, and the presiding judge decides which evidence is to be introduced at the trial.[205]

One important difference between German and American criminal procedure law lies in the absence of trial juries in Germany. Although lay judges sit together with professional judges in panels that try and decide nearly all non-petty criminal cases,[206] lay persons do not independently determine the verdict but instead deliberate and decide together with one or more professional judges. It is important to note that lay judges do not have access to the prosecutor's dossier.[207] This "mixed" composition of criminal courts has a substantial impact on German evidentiary law. Unlike in the U.S. and other jurisdictions that utilize juries as fact finders, German evidentiary rules are not concerned about shaping trial evidence in a way to avoid misleading jurors—there is always a professional judge available to explain to lay judges the relevance and possible pitfalls of evidence *in camera*.[208] German evidentiary law is therefore more liberal in admitting evidence, whereas U.S. law attempts to strictly control the input of evidence due to the influence it has on jury deliberations.

As the presentation of evidence at the trial is controlled by the presiding judge, the exclusion of evidence irrelevant to the resolution of the case rarely presents a problem. Judges typically do not wish to spend time on introducing evidence they regard as irrelevant. A relevance problem can arise, however, if one of the parties requests to have additional evidence introduced, either by the court or by themselves.[209] Typically the court

---

202. For a brief comparison of German and American procedural models, *see* THOMAS WEIGEND, *Modelle des Strafverfahrens: Deutschland und USA*, *in* VERWIRKLICHUNG UND BEWAHRUNG DES RECHTSSTAATS 31–45 (Eric Hilgendorf et al. eds., 2019).

203. STPO, §§ 238, para. 1, 244, para. 2, 245, para. 1, sentence 1.

204. *Id.* §§ 160, para. 1, 161, para. 1, 163, para. 1, 170, para. 1.

205. *Id.* §§ 170, para. 1, 199, 244, para. 2.

206. There exist differently composed mixed panels for hearing cases of lesser and greater severity. In some of these panels, lay judges have a majority. *See* Gerichtsverfassungsgesetz [GVG] [Courts Constitution Act] May 9, 1975, BGBl I at 1077, §§ 29, para. 1, sentence 1, 76, para. 1, last amended by July 7, 2021 (Ger.) [hereinafter GVG]. Yet, due to the fact that the members of the panel discuss all relevant issues of fact and law, it is a rare occurrence that lay judges outvote their professional colleagues.

207. *See* GVG § 76, para. 1; STPO § 199.

208. *See* John Langbein, *Mixed Court and Jury Court: Could the Continental Alternative Fill the American Need?,* 1981 AM. BAR FOUND. RSCH. J. 195, 198-202 (1981).

209. *Id.* §§ 244 para. 3–6, 245.

must comply with such requests, but the presiding judge can deny a request if the fact the evidence seeks to prove is irrelevant to the determination of the case[210] or if the evidence offered is not useful in proving the fact.[211] The latter condition has been found to exist, for example, when the defense offered the testimony of a parapsychologist[212] or the results of an *ex post facto* experiment[213] to demonstrate that the defendant was intoxicated at the time of the offense. As new forensic methods emerge, the questions of when to introduce expert evidence and what is considered expert evidence remain controversial.

German law has established several rules purporting to make judicial fact-finding trustworthy. One such rule restricts trial evidence to four types: Witness testimony, expert testimony, documentary evidence, and "proof by inspection" of objects that can be viewed or heard in court.[214] According to the so-called immediacy rule (*Unmittelbarkeitsprinzip*), testimony of a witness must not be replaced at the trial by the protocol of an earlier interrogation of the witness.[215] German law thus prefers live witness testimony over documentary evidence at trial. According to the majority view, this rule does not exclude hearsay testimony presented by a witness at trial.[216] However, the court's general obligation to determine the truth typically prompts the court to summon original witnesses where available. With regard to expert witnesses, the German Federal Court of Justice has ruled that their testimony must adhere to the standards of methodology applicable to their field of expertise[217] In reaching their judgment, the judges may rely on scientifically established findings of the expert, even if they cannot independently verify their validity; the principles and rules applied by the expert must, however, be generally accepted in the relevant scientific community.[218]

Once a verdict has been pronounced, the judges must write an extensive judgment in which they explain the evidentiary basis of their findings, detailing the evidence they found convincing and why.[219]The written judgment must contain an objective and consistent basis for the court's determination—mere assumptions or speculations will not

---

210. *Id.* § 244, para. 3, sentence 3.

211. STPO § 244, para. 3, sentences 2–3. It should be noted that the court can deny a request for taking evidence only if the proposed evidence is irrelevant on its face; the court is thus precluded from presuming what the proposed witness will testify and whether his testimony will appear to be credible. *See id.* § 244, para. 4, sentence 2.

212. Bundesgerichtshof [BGH] [Federal Court of Justice], Feb. 21, 1978, Neue Juristische Wochenschrift [NJW] 1207 (1978) (Ger.).

213. Bayerisches Oberstes Landesgericht [BAYOBLG] [Bavarian Higher Regional Court], Jan. 12, 1966, 12 JURISTISCHE RUNDSCHAU [JR] 227 (Ger.) (establishing radar functionality at the time a photograph is taken).

214. *See* Ulrich Eisenberg, BEWEISRECHT DER STPO [EVIDENTIARY LAW IN CRIMINAL PROCEDURE], marginal note 35 (2017) (Ger.).

215. STPO § 250; *but see id.* § 251, para. 1 (providing that a transcript from a prior interrogation can be introduced in lieu of in-person testimony where a witness has died or is not readily available, or if all parties agree).

216. *See* BERTRAM SCHMITT & MARCUS KÖHLER, *Strafprozessordnung: Gerichtsverfassungsgesetz, Nebengesetze und Ergänzende Bestimmungen*, *in* LUTZ MEYER-GOßNER & BERTRAM SCHMITT, STRAFPROZESSORDNUNG MIT GVG UND NEBENGESETZEN § 250 (66th ed. 2023) (Ger).

217. Bundesgerichtshof [BGH] [Federal Court of Justice] Dec. 17, 1998, 44 Entscheidungen des Bundesgerichtshofes in Strafsachen [BGHST] 308 (Ger.) (holding that the polygraph does not constitute a scientific method). *See also* 45 BGHST 164 (Ger). (promulgating rules for the scientific determination of witness credibility).

218. 44 Entscheidungen des Bundesgerichtshofes in Strafsachen [BGHST] 308 (Ger.); *see also* Klaus Miebach, MÜNCHENER KOMMENTAR ZUR STRAFPROZESSORDNUNG [COMMENTARY ON THE GERMAN CODE OF CRIMINAL PROCEDURE] § 261 marginal numbers 70–71 (Christoph Knauer et al. eds., 2016) (Ger.).

219. STPO § 267. The written judgment is authored by the professional members of the court; lay judges need not sign the document; *id.* § 275 para. 2.

suffice.[220] The obligation to provide extensive reasons encourages the court to draw rational conclusions on the accuracy and reliability of each piece of evidence. If there are apparent contradictions in the court's written judgment or if the evidentiary basis of its factual findings it's not sufficiently explained, the decision will be reversed on appeal.[221] German law thus provides for an effective *ex post* check on the trial court's decisions.

B. Defense Rights

Due to its basis in the inquisitorial tradition, the German Code of Criminal Procedure does not specifically provide for the right of the defendant to confront adverse witnesses at the trial. However, the right to confrontation, while traditionally associated with adversarial systems, has spread to European inquisitorial systems through article 6, paragraph 3 (d) of the European Convention on Human Rights ("ECHR") which provides that everyone charged with a criminal offense has the right to examine or have examined witnesses against him.[222] This language is derived from criminal procedure in adversarial systems, where there are witnesses for the prosecution and the defense. The European Court of Human Rights ("ECtHR") has interpreted the right of confrontation to mean that the defendant may examine the witnesses who actually observed the relevant occurrence; his right is not limited to the examination of hearsay witnesses.[223]

According to its language, article 6, paragraph 3(d) of the ECHR applies only to witnesses. However, the ECtHR interprets the term "witness" broadly to include expert witnesses,[224] victims, "and other persons testifying before the court."[225] Arguably, as suggested above with respect to the Confrontation Clause of the U.S. Constitution, an equivalent to the right to confrontation could be applied to device-generated data. If the findings and conclusions that are decisive in determining a defendant's guilt are generated independently (and in a partially unverifiable manner) by a device, the right of confrontation is not satisfied where a defendant is only given the opportunity to examine a programmer or an expert witness.

Article 6, paragraph 3(d) of the ECHR also contains a guarantee almost identical to the compulsory process clause of the Sixth Amendment to the U.S. Constitution:

> "Everyone charged with a criminal offence has the following minimum rights: (…) (d) … to obtain the attendance and examination of witnesses on his behalf under the same conditions as witnesses against him . . . ."

---

220. BGH, Feb. 7, 2012, 2012 Neue Zeitschrift für Strafrecht – Rechtsprechungsreport [New Journal of Criminal Law – Caselaw Reporter] 150 (Ger.).

221. Louisa Bartel, KARLSRUHER KOMMENTAR ZUR STRAFPROZESSORDNUNG § 267 marginal number 16 with references (Christoph Barthe and Jan Gericke, eds., 9th ed. 2023).

222. European Convention on Human Rights art. 6, ¶ 3(d), *opened for signature* Nov. 4, 1950, 213 U.N.T.S. 222 (entered into force Sept. 3, 1953).

223. Unterpertinger v. Austria, Eur. Ct. H.R., App. No. 9120/80, ¶¶ 31–32 (1986). *See also* Miranna Biral, *The Right to Examine or Have Examined Witnesses as a Minimum Right for a Fair Trial,* 22 EUR. J. CRIME, CRIM. L. & CRIM. JUST. 331 (2014); Tom Decaigny, *Inquisitorial and Adversarial Expert Examinations in the Case Law of the European Court of Human Rights,* 5 NEW J. EUR. CRIM. L. 149 (2014) (providing an overview of the relevant caselaw).

224. Khodorkovskiy v. Russia, Eur. Ct. H.R., App. No 11082/06 & 13772/05, ¶ 711 (2013); Ivanovski v. Former Yugoslav Republic of Macedonia, Eur. Ct. H.R., App. No. 10718/05, ¶ 56 (2014); *see also* Joëlle Vuille et al., *Scientific Evidence and the Right to a Fair Trial Under Article 6 ECHR,* 16 *L. PROBABILITY & RISK* 55 (2017).

225. Mirilashvili v. Russia, Eur. Ct. H.R., App. No. 6293/04, ¶ 158 (2008).

Although the German constitution does not provide for a defendant's right to present witnesses on his or her behalf, German procedural law accommodates the interests protected by the compulsory process clause. First, the trial court is obliged under the inquisitorial principle to summon all witnesses whose testimony may be relevant to the case; there is no distinction made between witnesses "for" or "against" the defendant.[226] Second, defendants have the right to summon witnesses,[227] and the trial court is bound to hear these witnesses unless the fact to which they are to testify is evident or has already been proven, the evidence proposed is not connected to the subject matter of the trial, or is not useful in resolving the case.[228]

C. Device Evidence Under German Law

Although German evidentiary rules are less rigid than those found in the U.S., they nevertheless place a strong emphasis on establishing a threshold of trustworthiness for evidence introduced at trial. With respect to evidence generated by devices, the first (and essential) question is whether the evidence fits into one of the four permissible types of trial evidence, i.e., witness or expert witness testimony, written documents or physical evidence.[229] Since device evidence does not typically take the form of a written document, the remaining potential evidence types include witness testimony, inspection of physical evidence, and expert testimony.

i. Witness Testimony

If, as in the Swiss case described above, drowsiness warnings recorded by a driver's assistance system are to be offered as proof of the driver's negligence, one might think of treating that device's observation as witness testimony due to its recording of past facts and conditions.[230] However, as German procedural law stands today, only human beings can be witnesses, because devices are not capable of making verbal statements or of answering questions posed by a judge or the parties at trial.[231] While much research has been focused on developing "explainable" AI, progress to date has not been such that "device witnesses" capable of explaining their assessments can be expected to walk into courtrooms anytime soon.[232]

ii. Proof by Inspection

If an entity cannot testify verbally, German law provides for the introduction of physical evidence as "proof by inspection" *(Augenscheinsbeweis)*.[233] Under this option, the judges visually or aurally inspect objects in court with the parties present.[234] Proof by inspection is employed, for instance, when a photo is presented that has been taken by a radar gun

226. *Cf.* Eisenberg, *supra* note 216, at 138.
227. STPO § 220.
228. STPO § 245 (2).
229. *See* Eisenberg, *supra* note 216.
230. For a definition of witness, *see* Eisenberg, *supra* note 216, at 1000.
231. STPO § 57, 59, 68–68a (governing the instruction of the witness, the possibility of placing the witness under oath, the examination as to the witness's identity, and the limitation of examination to protect the privacy of the witness).
232. If robots were to function as witnesses, the question of applicability of testimonial privileges would likely arise. In Germany the law currently only addresses privileges for close relatives and members of certain professions, *see* STPO §§ 52, 53.
233. *See* STPO § 86.
234. *Id.*

and is offered to prove a violation of a speed limit and to establish the driver's identity. With respect to device evidence, the question becomes whether it is possible for the device's "findings" to be converted through a standardized and robust method into visual objects that can be observed in a courtroom akin to radar photos.[235] This would not only require standardization of the methods of data generation[236] and storage,[237] but also of the techniques for visualization.[238] If such steps can be taken, warnings issued by a drowsiness monitor, for example, could potentially be presented as visual documentation that the judges and trial parties could view and discuss.

iii. Expert Evidence

While German law expects judges to comprehend observations provided by human witnesses and objects presented for inspection, the court must appoint experts where the judges lack the expertise to properly evaluate evidence.[239]

Thus, if a device cannot take the stand as a witness and its data cannot be brought to a courtroom for inspection, the next best option may be to request an expert to evaluate the information before trial and subsequently testify about it. German courts routinely hear expert evidence, for example, on data stored in a car's event data recorder just before an accident.[240] In contrast to evidence introduced as "proof by inspection," evidence introduced via expert witness testimony allows the court to address a human being who assumes responsibility for the interpretation of the data generated by the device and who can respond to case-specific inquiries. Additionally, the defense can question experts and may also request that additional experts be appointed or provide its own experts.[241]

Under German law, an expert testifies on matters that are not accessible or comprehensible to lay persons, including the judges.[242] With regard to intelligent devices, expertise would be necessary to understand how the system works, how the data is generated, stored and reproduced, and whether the data is accurate.[243] An expert could also interpret device-generated data and opine on the probability that it supports the facts at issue.[244] Expert testimony may thus be the preferred way of introducing device-generated data at trial; but, even the best expert cannot make device evidence more accurate and reliable than it is. Experts can, however, testify as to whether a device's data can be validated.

---

235. *Id.*
236. The EU Commission's proposed Artificial Intelligence Act might be helpful in setting such standards, but not without documentation obligations. *See Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonized Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts*, COM (2021) 206 final (Apr. 21, 2021).
237. For automated cars, relevant data would be stored in the car's Data Storage System for Automated Driving (DSSAD), for more details see Gless et al., *supra* note 128, at 288, 290.
238. Gless et al., *supra* note 128, at 289–90.
239. *Cf.* STPO § 244.
240. *Cf.* NAT'L HIGHWAY TRAFFIC SAFETY ADMIN. [NHTSA], EVENT DATA RECORDER, https://www.nhtsa.gov/research-data/event-data-recorder (last visited Nov. 6, 2023).
241. STPO § 244 para. 4.
242. *See* CLAUS ROXIN AND BERND SCHÜNEMANN, STRAFVERFAHRENSRECHT § 243 (30th ed. 2022).
243. Andreas Winkelmann, *„Einzelraser" nach §315 d Abs. 1 Nr. 3 StGB und der Nachweis durch digitale Fahrzeugdaten*, 19 DEUTSCHES AUTORECHT [German Automobile Law] 1, 2–6 (2023).
244. Such as: "The human driver was too tired to drive properly" versus "The driver assistance system malfunctioned." *Cf.* Robert Cook et al., *A Hierarchy of Propositions: Deciding Which Level to Address in Casework*, 38 SCI. & JUST. 231–32 (1998).

In accordance with the German rule of immediacy (*Unmittelbarkeitsprinzip*), as well as the principle of orality, expert witnesses most often testify in open court where they are subject to questioning by judges and the trial parties.[245] But experiments and tests can be conducted by an expert before the trial and the expert may then report at trial about the methods used and the conclusions drawn.[246] Given that German procedural law does not specifically prohibit the introduction of hearsay evidence, [247] the expert witness is free to use information provided by others. Moreover, in less contested cases the court may elect to dispense with the personal appearance of a sworn expert and instead have his or her written report read aloud in court.[248]

iv. Testing Accuracy and Reliability

Inquisitorial proceedings in Germany place the responsibility for testing and determining the accuracy of evidence exclusively on the trial court, which makes a decision based on its appraisal of the totality of the evidence presented at the trial.[249] In contrast to some other European jurisdictions, there is no "investigating magistrate" tasked with assessing the evidence before trial.

D. The Defense's Right to Evaluate Evidence

The defense can play a significant role in the process of evaluating the accuracy and reliability of trial evidence. Under German procedural law, the defendant is accorded early access to information in the prosecutor's case file that forms the basis of the accusation. The defense lawyer may demand to inspect the prosecution dossier at the latest after the conclusion of the investigation.[250] This is in line with Art. 6 (1) ECHR, which guarantees the defendant the right to view incriminating evidence to be presented by the prosecution.[251] At the trial, evidence is introduced in the presence of the parties. They may question each witness[252] in accordance with the defense's confrontation rights as provided in Art. 6 (3) (d) ECHR. Parties may also request the court to take additional evidence,[253]

---

245. *See* Helmut Kreicker, MÜNCHENER KOMMENTAR ZUR STRAFPROZESSORDNUNG [Commentary on the German Code of Criminal Procedure] § 250 (Christoph Knauer et al. eds., 2016) (Ger.).

246. *Cf.* BERTRAM SCHMITT & MARCUS KÖHLER, *Strafprozessordnung: Gerichtsverfassungsgesetz, Nebengesetze und Ergänzende Bestimmungen*, *in* LUTZ MEYER-GOßNER & BERTRAM SCHMITT, STRAFPROZESSORDNUNG MIT GVG UND NEBENGESETZEN §79 (66th ed. 2023) (Ger).

247. *See* STPO § 250.

248. STPO § 256.

249. STPO § 261.

250. The defense lawyer may request to inspect the prosecutor's file at any time, but until the conclusion of the investigation the prosecutor may deny inspection if sharing the information may impair the investigation, for example, because the defendant may abuse the information for influencing witnesses. STPO § 147 (2).

251. *See* Brandstetter v. Austria, 211 Eur. Ct. H.R. 4, 22 (1991) (interpreting the principle of equality of arms as one aspect of a fair trial); *see* Dowsett v. United Kingdom, App. No. 39482/98 Eur. Ct. H. R. 261, 275, 277 (2003) (finding that the defendant's rights were violated when the court ordered an essential piece of evidence to be destroyed before the defense lawyer could inspect it). *See also* Papageorgiou v. Greece, App. No. 59506/00 Eur. Ct. H.R. 245, 252 (2003) (holding "[t]he right to an adversarial trial means, in a criminal case, that both prosecution and defence must be given the opportunity to have knowledge of and comment on the observations filed and the evidence adduced by the other party . . . Article 6 § 1 requires that the prosecution authorities should disclose to the defence all material evidence in their possession for or against the accused."); Baumet v. France, App. No. 56802/00, Eur. Ct. H. R. 1, 9–13 (2007) (finding a violation of Article 6 § 1 where a prosecutor submits documents to the court without informing the defendant).

252. STPO § 240 para. 2.

253. STPO § 244 para. 3.

in which case it would evaluate the relevance of the proposed evidence and deny the request only where it was deemed irrelevant, redundant, not useful, or unattainable.[254]

It remains unclear in what way these defense rights are to be applied to device evidence. In the digital age, the traditional defense right to inspect the prosecutor's dossier may be insufficient where incriminating evidence is delivered by non-human devices. Defendants have an interest in not only receiving the evaluative data the device has produced, but in learning how and on what basis the device came to its conclusions. This leads to the question of whether the defendant can derive from the right to a fair trial a right to have a device's "decision-making processes" disclosed in a verifiable manner.

German courts have addressed this problem. In a 2020 case involving digitized radar guns, the Federal Constitutional Court[255] held that the right to a fair trial in principle includes a right to obtain access to all relevant raw and/or measurement data that have been stored for the purpose of the investigation, even if they were not included in the case file.[256] The Court has recognized a "right to raw data" based on Article 2 in conjunction with Article 20 of the German Basic Law[257] and emphasized the importance of being able to trace the machine's data processing operations.[258] Even before the 2020 landmark decision, some courts had argued that defendants must be able to investigate whether there exist any doubts about the viability of the accusation; if they cannot do so, the factual basis of the conviction would ultimately be shielded from meaningful verification.[259] Under this case law, the driver in the case of a drowsiness alert would have to be granted access to the raw measurement data, the algorithms, and the source code that determined the triggering of the device's activity. As our simulated BMW cross-examination demonstrated, such data may be of limited use. If German courts strictly applied their rulings to evidence produced by devices with an evaluative dimension, they would probably have to negate its admissibility since the defendant is unable to exercise his right to information if the way in which the information was generated remains inscrutable.

IV. CORE PROBLEMS AND SOLUTIONS

A. The Core Problem of Device Evidence

Let us now compare the approaches of the American and German systems of evidence. Both legal systems have in common that they are profoundly humanistic. For centuries they have relied on data from human beings that is evaluated by human beings.

---

254. *Id.*

255. The Federal Constitutional Court (*Bundesverfassungsgericht*) can, upon an individual's complaint, review any German court's final judgment for possible violations of the complainant's constitutional rights; if a violation has been found, the Federal Constitutional Court will overturn the impugned judgment. See *Basic Law of the Federal Republic of Germany Art. 93 subsec. 1 no. 4a* (*Grundgesetz für die Bundesrepublik Deutschland*), FEDERAL MINISTRY OF JUSTICE, https://www.gesetze-im-internet.de/englisch_gg/ (Eng).

256. Bundesverfassungsgericht [BVerfGE] [Federal Constitutional Court], 2 BvR 1616/18, Nov.12, 2020, (Ger.)
https://www.bundesverfassungsgericht.de/SharedDocs/Entscheidungen/DE/2020/11/rk20201112_2bvr161618.html.

257. GG, Art. 2 subsec. 1 and Art. 20 subsec. 3.

258. Bayerisches Oberstes Landesgericht [BayObLG] [Bavarian Higher Regional Court], Dec. 9, 2019, 202 [ObOWi] 1955/19 (Ger.), https://www.gesetze-bayern.de/Content/Document/Y-300-Z-BECKRS-B-2019-N-31165?hl=true (reversing prior decisions denying an obligation to disclose such data due to the assumption that calibrated and regularly monitored devices produce valid findings).

259. Oberlandesgericht Saarbrücken [Higher Regional Court of Saarland] Sep. 3, 2019, Ss Rs 34/2019 [43/19 OWi] (Ger.), https://www.burhoff.de/asp_weitere_beschluesse/inhalte/5294.htm.

If device evidence is to be admitted in court, judges are therefore expected to largely rely on expert testimony for determining its accuracy.[260] Human experts must analyze the data in the context of the circumstances in which it originated[261] and reach conclusions based on their expertise. The difficulty of determining the accuracy of device evidence is proportional to the sophistication of the device.[262] With each layer of autonomy added, the retracing of a device's assessment becomes more difficult.[263]

The challenges are especially acute if a device draws its own conclusions from the information it has gathered without recording all conditions for its conclusion.[264] In that instance, not all outcomes may be traceable to humans, even if the underlying raw data, measurement data, source code and algorithmic processing are available.[265] Returning to the case of the driver drowsiness warning system, we know that following the accident it was determined, by reading out the car's data storage system,[266] that the driving assistant warned the human driver. But, in the absence of additional data to provide context, it is not possible to determine why the warning was issued, whether the assessment carried out by the device was based on actual signs of fatigue, a misinterpretation, or a processing error in the data measured. As our BMW cross-examination suggests, there may have been a viable justification for the human driver's action that the driving assistant was not programmed to recognize, let alone understand. If the court simply relies on the findings of a drowsiness warning system, it relies, in the final analysis, irrationally—one could even say blindly—on the assumption that the device recorded the relevant data correctly and drew accurate conclusions. Making such an assumption would violate the requirement of a rational explanation of the verdict, i.e., a sound explanation that is transparent in its reasoning. It is another question, however, whether a drowsiness detection system's alert could be used as circumstantial evidence to show how events transpired.

This problem may have different consequences in inquisitorial and adversarial procedural systems. From the perspective of an inquisitorial-type system, totally rejecting device-generated data would foreclose a potentially important source of information for the court and thus might increase the risk of miscarriages of justice.[267] There would be considerable pressure to accept a device's assessment of the performance of humans as evidence. This is especially true because device-generated evidence may be more reliable than the testimony of human witnesses.

---

260. For more detailed inquiries into this problem, *see* Edward J. Imwinkelried, *Improving the Presentation of Expert Testimony to the Trier of Fact: An Epistemological Insight in Search of an Evidentiary Theory*, 52 ARIZ. ST. L. J. 49, 57–59 (2020); Joëlle Vuille & Franco Taroni, *Measuring Uncertainty in Forensic Science,* IEEE INSTRUMENTATION & MEASUREMENT MAG. 8 (2021); Murphy, *supra* note 51; Steven P. Lund & Hari Iyer, *Likelihood Ratio as Weight of Forensic Evidence: A Closer Look,* 122 J. RES. NATL. INST. STAND. TECHNOL. 1 (2017).

261. Alex Biedermann & Joëlle Vuille, *Digital Evidence, 'Absence' of Data and Ambiguous Patterns of Reasoning*, 16 DIGIT. INVESTIGATION, 86, 90 (2016); Sandra Wachter & Brent Mittelstadt, *A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI*, 2 COLUM. BUS. L. REV. 494, 510–11 (2019).

262. *See* Gless, *supra* note 59, at 211–12.

263. *Id.* at 211.

264. *Id.*

265. *See generally* Edward J. Imwinkelried, *Computer Source Code: A Source of the Growing Controversy Over the Reliability of Automated Forensic Techniques*, 66 DEPAUL L. REV. 97 (2016).

266. For more information on the Data Storage System for Automated Driving (DSSAD), *see* Gless et al., *supra* note 128, at 289–93.

267. In 2013, the Federal Constitutional Court declared, "[i]t is the central concern of criminal proceedings to establish the true facts of a case without which it is impossible to implement the substantive principle of individual guilt." BVerfG [Federal Constitutional Court], 2 BvR 2628/10, Mar. 19, marginal number 56.

Traditionally, the inquisitorial system does not emphasize adversarial vetting mechanisms but places its trust in the judges' professional experience in assessing evidence.[268] The belief in the judges' competence in reliable fact-finding seems to persist even when courts are faced with new developments, like the emergence of device evidence. The adversarial system, by contrast, is inherently more critical of the reliability of evidence introduced by the parties. There is still a strong belief in the effectiveness of antagonistic confrontation, including cross-examination as "the greatest legal engine ever invented for the discovery of truth."[269] But some scholars have observed a decrease in effectiveness of traditional safeguards within the adversarial process,[270] and, in particular, for ensuring the integrity of complex technology-based forensic evidence. If the provider of information is not a human being, but a device incapable of telling lies and unable to explain its "thought" processes, cross-examination cannot fulfill its function. Consequently, the need to scrutinize the reliability of data generating devices outside rather than in the courtroom may arise.[271]

There are two potential solutions to the problem of ensuring the accuracy and reliability of device data, a technological answer and a procedural answer.

B. A Technological Solution

As the collecting, processing, and evaluation of data by devices often are not fully comprehensible to humans, confidence in the accuracy of their observations and assessments could be bolstered by means of technological standardization[272] and certification,[273] as well as continuous device inspection and calibration. Some scholars have proposed formal reliability validation frameworks and taxonomies for the assessment of digital forensics in criminal cases, based on validation criteria and validation testing techniques.[274] These frameworks could help judges and defense lawyers to better understand reliability issues and to efficiently test forensic reports.[275] However, the most promising path may lead toward special types of tools built to assist human assessment of evidence[276] or even to completely take over the vetting of device evidence and to delivering authenticity certificates.[277]

---

268. *See supra* note 209 and accompanying text. Even when lay judges are involved, they do not engage in fact-finding by themselves but do so together with the professional judges.

269. California v. Green, 399 U.S. 149, 158 (1970) (quoting 5 WIGMORE § 1367).

270. *See, e.g.*, Ortman, *supra* note 55; Rebecca Steele*, Equalizing Access to Evidence: Criminal Defendants and the Stored Communications Act*, 131 YALE L. J. 1584 (2022); Sela Brown, *Brady in the Plea Era: How U.S. v. Ruiz Should Be Reconstrued in Light of Missouri v. Frye and Lafler v. Cooper*, 27 BERKELEY J. CRIM. L. 1 (2022).

271. *See, e.g.*, Murphy, *supra* note 51.

272. Such standards are set for driving assistants in vehicles, *cf.* Regulation (EU) 2019/2144, Art. 4.

273. With regard to AI this approach is proffered in the draft EU-Act as well with regard to "high-risk" artificial intelligence systems *cf*. Art. 12, 44 of the draft AI-Act. It relies on established of forensic evidence.

274. Radina & Katrin Franke, *Reliability Validation Enabling Framework (RVEF) for Digital Forensics in Criminal Investigations*, 45 FORENSIC SCI. INT'L: DIGIT. INVESTIGATION 301554 (2023); Rune Nordvik et al., *Reliability Validation for File System Interpretation*, 37 FORENSIC SCI. INT'L: DIGIT. INVESTIGATION 301174 (2021); Julia Simon-Kerr, *Credibility in an Age of Algorithms*, 74 RUTGERS U.L. REV. 111 (2021).

275. Stoykova & Franke, *supra* note 290.

276. *Cf.* Stefania Costantini et al., *Digital Forensics and Investigations Meet Artificial Intelligence,* 86 ANNALS OF MATHEMATICS AND A.I. 193 (2019); Nick L. Petroni Jr.a , Aaron Waltersb, Timothy Frasera & William A. Arbaugh, *FATKit: A Framework for the Extraction and Analysis of Digital Forensic Data from Volatile System Memory*, 3 DIGIT. INVESTIGATION 197 (2006).

277. *Cf.* Cosimo Anglano et al., *The Android Forensics Automator (AnForA): A tool for the Automated Forensic Analysis of Android Applications*, 88 COMPUTS. & SEC. 1 (2020); Aaron Jarrett & Kim-Kwang Raymond Choo, *The Impact of Automation and Artificial Intelligence on Digital Forensics*, 3 WIREs FORENSIC SCI. 1 (2020).

Whichever path is chosen, the first prerequisite for a technological solution would be the development of standards for the relevant categories of data[278] and for their subsequent retrievability.[279] Interdisciplinary research teams can develop testing processes for assessing the technical reliability of devices and for determining the accuracy of the data they generate.[280] In the development of such a device, trade-offs will be necessary. If the aim is a fully automated validated algorithmic solution, the complexity of the required algorithms will be very high and their explain-ability very low.[281] Such an approach can however be justified if such tools enhance the overall quality of the criminal process. If such tools are used by many law enforcement agencies, flaws could be detected by an input-output check of the overall results. One progenitor of such kind of standardisation in forensic software could be the Crash Data Retrieval ("CDR") tool designed to access and retrieve data stored in an Event Data Recorder. CDRs are now standard in cars manufactured in the United States.[282] On the basis of such a validation process, it would then be possible to determine whether, for example, a driver assistance system can be accepted as ordinarily error-free. Validation could be based on test datasets that would permit confirmation of results with a certain degree of probability.[283] Using an AI device to access and process the relevant data and its handling would eclipse inevitable human limitations on managing heterogeneous data.

If such a process has been chosen, statutes and regulations could stipulate that only certified machine learning devices may be used or that only their data are admissible in court.

Unfortunately, even certifying and checking systems that produce device evidence may not completely solve the problem of the limited ability to explain a device's assessment. Even if an assistance system has been tested and certified prior to its entry into the market, there is no guarantee that it will never draw an incorrect conclusion.[284] Thus, it is possible that a car will steer to the left side of the road even when it should not do so, and a human jury will be unable to determine why the device erred. Take, for example, a defendant who challenges the accuracy of her car's driver assistance system that assessed her steering movements as erratic by pointing out that she was driving on a road without markings and the car mistook the left-hand roadside for the middle road marking. While the certification process may demonstrate the system's general reliability,

---

278. *See* Paul W. Grimm et al., *Authenticating Digital Evidence*, 69 BAYLOR L. REV. 17, 41 (2017).

279. *Cf. e.g.*, Agreement Concerning the Adoption of Harmonized Technical United Nations Regulations for Wheeled Vehicles, Equipment and Parts, E/ECE/TRANS/505/Rev.3/Add.156 of March 4, 2021, no. 8 'Data Storage System for Automated Systems'; reading out the data will be possible by using On-Board Diagnostics Port, 2nd generation (OBD II port), launched in 1996, for further information, *see UNECE is Driving Progress on Autonomous Vehicles,* UNECE, https://unece.org/automated-driving (last visited Jan. 9, 2024); *see also* Stoykova & Franke, *supra* note 290; Nordvik et al., *supra* note 290.

280. *See e.g.*, Anglano et al., *supra* note 293 for automated forensic analysis of android applications (like text messaging or photo posting, GPS coordinates).

281. *Cf.* Buhrmester et al., *supra* note 105, at 984.

282. *See* NHTSA Event Data Recorders Rules, 49 C.F.R. pt. 563, https://www.law.cornell.edu/cfr/text/49/part-563; Jeremy S. Daily, Nathan Singleton, Elizabeth Downing & Gavin W. Manes, *The Forensics Aspects of Event Data Recorder*s, 3 J. DIGIT. FORENSICS, SEC. & L. 29 (2008); Le-Khac et al., *supra* note 13, at 503; Gless et al., *supra* note 128, at 287–88.

283. *Cf.* Jason Brownlee, *What is the Difference Between Test and Validation Datasets?* (Aug. 14, 2020), https://machinelearningmastery.com/difference-test-validation-datasets.

284. As explained above, incorrect training data can lead to false results, *cf.* Buhrmester et al., *supra* note 105, at 966–67.

information on how the system draws a certain conclusion in a specific real-time situation cannot be simulated in advance. There is also always the additional risk of manipulation of the device by a third party via an unknown security leak.

Our analysis of the problems inherent in device evidence reflects current and near-future technology. Ultimately, technology itself may be able to resolve our concerns.[285] A promising avenue might be the use of AI to check and possibly verify devices' assessments.[286] Inspired by the English term for a counter activity to hostile intelligence activities, we call such (future) verification software "Artificial Counter-Intelligence" ("ACI"). The basic idea of ACI is to check the operational reliability of a device—regardless of the underlying technology—without having to rely on human input. This might be an outgrowth of the current use of adversarial AI to train AI systems to avoid inaccurate data.

ACI could provide general information about the functionality of a device, thus enabling the factfinder to assess the reliability of its output.[287] To do this, ACI could run a predefined simulation of the raw measurement data stored, for instance, in a vehicle's DSSAD. Rule-based systems with their pre-programmed data processing procedures could be verified in this way through comparatively transparent means; case analysis could be used to make sure that the tool performed with a sufficient degree of accuracy. ACI could be used to check the reliability of Type 2 devices, which, as we have seen above, may have undergone complex machine-learning processes that created faults and cannot be traced by human beings when they come in a black box.[288] ACI is not science fiction; rather it pushes the envelope on various initiatives for trustworthy AI[289] and existing AI tools for specific forensic applications.[290]

The idea of using ACI to "vet" the data generated by devices raises a number of fundamental questions: What is the relationship between data accuracy and "truth"? And how can "truth" be operationalized?[291] Can the same rules be used to ensure the trustworthiness of evidence from human witnesses and from devices, or is it necessary to develop entirely new rules especially for devices, perhaps via adherence to International Organization for Standardization (ISO)[292] or other technical standards? If ACI were to be used to assist in the admission of device evidence in court, close cooperation between lawyers and experts in AI technology would be necessary. Lawyers must work out the normative requirements for the use of device evidence. AI specialists must then link these legal principles to heuristic decisions and patterns of machine learning by devices. Minimal standards of reliability of both devices and ACI for civil and criminal proceedings may then be possible.

---

285. Anglano et al., *supra* note 293.

286. For a proposal to use automated forensic tools in the specific case of android applications on smartphones: *see* Anglano et al., *supra* note 293.

287. *Cf. id.*

288. Buhrmester et al., *supra* note 105.

289. Luciano *Floridi, Establishing the Rules for Building Trustworthy AI*, 1 *NATURE MACH. INTEL.* 261 (*2019).*

290. *Cf.* Anglano et al., *supra* note 293.

291. *Cf.* Roth, *supra* note 8, at 2005; Christian Chessman, *A "Source" of Error: Computer Code, Criminal Defendants, and the Constitution,* 105 CAL. L. REV. 179, 188–89 (2017).

292. The International Organization for Standardization (ISO) is an international standard-setting body composed of representatives from various national standards organizations <www.iso.org>. Of relevance for the area discussed in this article are ISO/IEC 27043:2015 (information technology, security techniques, incident investigation principles and processes), ISO/IEC 27037:2012 (guidelines for identification, collection, acquisition and preservation of digital evidence), and ISO/IEC 27040 (storage security) (accessible at: www.iso.org/standard/44407.html).

For the use of ACI in court proceedings, there is the additional problem of how to apply evidence rules to the results of pretrial technological inquiry. Both relevance and authentication in the Federal Rules of Evidence are humanistic and cannot easily be transferred to technologies. However, the Federal Rules of Evidence have increasingly approved the use of text certifications from persons such as document custodians.[293]

Finally, the procedural connection between device evidence and its verification by ACI must be established. Here, at least two approaches are possible. One possibility would be to admit device-generated data as evidence only after the data has been checked by ACI and this verification process has indicated a certain – legally defined – level of reliability, in which case it would be up to the court to evaluate the evidence and to decide whether and to what extent to base its judgment on it. A second, more device-friendly approach would be to make device evidence generally admissible, but to grant the court, as well as the parties, the right to demand an ACI check. The latter option would have the advantage of speeding up the proceedings and reducing costs, which would make this option attractive at least in cases in which no serious objections to the device evidence were raised by either side.

C. A Procedural Solution

For better or worse, we do not at present have ACI as a generally employable tool to vet device evidence. Therefore, device evidence requires experts to conduct tests, present data, and relate their conclusions to courts. Given the partisan nature of expert testimony in the United States and the financial burdens on criminal defendants who need the assistance of experts, this places criminal defendants at a great, perhaps insurmountable disadvantage, which raises due process, confrontation and compulsory process issues.[294] These issues do not admit of easy solutions within the current United States legal culture. Therefore, judges and lawmakers may be well advised to draw on procedural solutions already in place in other countries.

German law, as we have seen, is based upon central control of the trial process by professional judges. The judges have the fundamental obligation to ensure that the evidence they introduce is reliable. Judges will appoint experts for the court whenever they think that the court lacks expertise on a matter relevant to its decision.[295] Court-appointed experts have access to the evidence and can test the theories of the prosecution and the defense about the events.[296] They will receive the requisite fee (according to a schedule determined by statute) and their necessary expenses from the state.[297] If a defendant is convicted, he or she is liable to pay the amount of the fee as part of court costs.[298] But if the defendant is indigent and in prison, it will often be difficult to enforce this obligation, so that the fee and expenses of the expert will ultimately be borne by the state. If the defendant hires their own expert, he or she has to pay the expert's fees and reasonable

293. Fed. R. Evid. 902.
294. Murphy, *supra* note 51, at 672.
295. StPO §73.
296. StPO § 80.
297. StPO § 71.
298. StPO § 464a, 465.

expenses in advance, or the expert can refuse to appear in court.[299] The German system thus still puts indigent defendants at a disadvantage with regard to hiring experts. However, the court's duty to appoint experts whenever necessary for discovering the "truth" makes it more likely that neutral experts will appear in German trials. A party can even request the recusal of an expert if the expert's conduct or prior announcements have given rise to doubts about his objectivity.[300] Forensic institutes are often accredited by the state to ensure their quality.

In the United States, the Supreme Court's holding in *Daubert* emphasized the importance of judges determining the validity of technological evidence but left it to adversarial experts to deal with proof.[301] Although Federal Rule of Evidence 706 allows federal judges to appoint non-partisan experts, it is uncommon in practice as it flies in the face of the adversarial process.[302] Yet, scholars have maintained that the discovery of scientific evidence can be "a game of cat and mouse" when "high-tech evidence" ought to be subject to extensive pretrial disclosures and depositions.[303] There have consequently been demands for a new regime that could help under-resourced defendants make use of government employed laboratory personnel and nurture a more neutral culture of science rather than a highly partisan atmosphere that damages forensic reputation.[304]

We note that we are unaware of any empirical data that would prove that German practice is superior to that of the United States in device evidence cases. But the combination of partisanship and financial inequity in the United States certainly suggests that we can and must do better than traditional practice.

V. CONCLUSION

The amount and importance of device evidence in criminal cases is bound to increase with the digital turn that has led to profound changes in transportation, medicine, and other important areas of human life. This development has been accompanied by an extensive monitoring of the human-robot-interaction, necessary to ensure safety (as in driving automation). While device evidence gains impetus, it presents significant legal questions in adversarial as well as inquisitorial legal systems. This is true especially when the devices that generate data proffered as evidence in a criminal case were not developed with criminal evidence law in mind. Criminal courts therefore face the question of whether to admit various kinds of device data offered as evidence.

The most tempting approach to this issue is to take the well-trodden path and apply the traditional rules on admissibility of evidence to device evidence. The courts have long dealt with evidence that raises substantial reliability and accuracy issues. Whether hearsay evidence or novel and cutting-edge scientific evidence, American courts have formulated rules and procedures for dealing with them. And American courts have managed to operate successfully with general public acceptance. But public acceptance might not continue if the public concludes that cases are determined by unreliable machines and/or

299. STPO § 220 para. (2).
300. STPO § 74; *see also* STPO § 24 (discussing the recusal of judges).
301. Daubert v. Merrell Dow Pharmaceuticals, 509 U.S. 579 (1993).
302. FED. R. EVID. 706 Advisory Committee's note to 1987 amendment.
303. Murphy, *supra* note 51, at 650–51.
304. *Id.* at 650.

that defendants face an unfair disadvantage because they lack the means to hire experts whose contribution is crucial to the selection and assessment of device evidence.

That concern leads us to recommend an alternative approach: American courts should adopt a rule of judicial responsibility for expert testimony in device data cases to ensure the integrity of all evidence proffered in criminal cases.

Given the nature of the adversary system, the ability of all parties to call expert witnesses moderates the factfinders' limited knowledge. Unfortunately, the impact of limited funds, especially in criminal cases, and the partisan character of expert testimony common in American trials may obfuscate the technological facts critical to the decision of a case. Device evidence demands a sort of collective inquiry instead of the individualized approach that heavily depends upon the skill of counsel and in-court confrontation rather than out-of-court expertise.[305] Accordingly, and in the spirit of the Supreme Court's decision in *Daubert v. Merrell Dow Pharmaceuticals*,[306] we recommend that Federal Rule of Evidence 706 be revised to encourage the appointment of experts by the court and to establish a procedure whereby the judge calls and neutrally examines the court-appointed expert when there are questions about the validity of scientific or technological evidence. Such examination would be followed by examination by the parties and then testimony of expert witnesses retained by the parties.

While this procedural recommendation would help resolve the concerns addressed by this article, given the current state of technological development, it may not be enough to ensure adequate factual determination in criminal cases that are heavily device dependent. Given the need for both factual accuracy and public acceptance of verdicts, we propose a new evidentiary rule based loosely on Federal Rule of Evidence 403, which declares in relevant part:

> The court may exclude relevant evidence if its probative value is substantially outweighed by a danger of . . . unfair prejudice.

The new evidentiary rule, Rule 403A,[307] would declare:

> *The court shall exclude otherwise relevant evidence when its source is data from a technological source the reliability and accuracy of which cannot reasonably be determined.*

In short, if a future BMW 7500i, Vehicle Number 12778899-x, belonging to AI-driven (Type 3) devices, is called to "testify," the judge could first examine the testimony of the court-appointed expert as to the reliability and accuracy of the BMW's proposed testimony and if its underpinnings are inadequate could simply rule: "Pursuant to Rule 403A, I hold that the proposed 'testimony' is inadmissible." Perhaps then the future BMW 7500i, Vehicle Number 12778899-x, and its technologists will retire from the courtroom vowing to do a better job of explaining its operation.

The Germany system does not escape our concerns, although it has a better legal framework with which to deal with them. Our conclusions can be transferred, *mutatis mutandis*, to German evidence law. While the appointment of neutral experts by the court

---

305. *Id.* at 672.
306. 509 U.S. 579 (1993).
307. This could also be denominated Rule 702A, as Rule 702 and subsequent rules deal with expert testimony typically related to scientific, medical, and technological evidence. *See* FED. R. EVID. 702.

is already provided for in the Code of Criminal Procedure,[308] introduction of evidence from a device based on inscrutable machine learning presents a problem even under the "all in" German approach to admitting evidence. The presiding judge can decline to admit evidence offered by a party if that evidence is deemed not of use,[309] but the court itself should also not introduce unusable evidence. And, as we have seen, actions taken by Type 3 devices cannot rationally be linked to human activity.[310] Because German law places less emphasis on input control, its structure may be more lenient toward Type 3 evidence in the future. However, admissibility of such evidence as "useful" would require that the members of the trial court, assisted by expert testimony, are able to understand, at least in a general way, the processes that take place in the operation of the device and the limitations of its output. If that were possible, German courts could be in the position to evaluate the weight given to the data produced by such a device.

In sum, we can say that measuring data from Type 1 devices can be introduced as evidence, provided that the reliability of the device's operation has been established. Data created by Type 2 devices can be used as evidence with certain precautions in place, typically through expert testimony, but their reliability could also potentially be enhanced through "Artificial Counter-Intelligence" devices. Data from Type 3 AI devices should not be admitted as long as their accuracy cannot be validated.

We are well aware that what we propose here would require significant changes to American evidence law, changes based on an unusual openness for legal solutions devised abroad. But when faced with the new phenomenon of device evidence, judges and lawyers need to be inventive and courageous.

308. STPO §73.
309. STPO § 244, para. 3.
310. Lake et al., *supra* note 151.