Stanford Law Review

Volume 76 April 2024

NOTE

Meaningful Machine Confrontation

Benjamin Welton*

Abstract. Machine-generated evidence is now ubiquitous in criminal trials, and more sophisticated forms of inculpatory evidence are on the way. Courts have almost universally held that the Confrontation Clause does not give criminal defendants a constitutional right to confront machine-generated evidence, except in narrow cases where the evidence also contains testimonial statements made by a human operator. Several scholars have countered that the Confrontation Clause should be read more broadly to consider machine accusers as "witnesses" that trigger confrontation rights. While cross-examination has been the traditional confrontation right in the American legal system, there are machine analogs such as source code disclosure, broadened discovery, machine access, and expanded live testimony that could be afforded to criminal defendants facing machine-generated evidence.

This Note presumes that the right to machine confrontation should exist and focuses on which of the proposed alternatives constitutes meaningful machine confrontation in light of the technology at issue. Existing calls for machine confrontation propose blanket solutions without fully considering the practical concerns judges invoke, the realities defense attorneys face, or the types of technology involved. This Note aims to fill this gap with three contributions. First, it provides a comprehensive view of the technologies entering criminal trials, including those sophisticated tools that will one day dominate discussion. Second, it proposes a taxonomy of the technological characteristics underlying these tools, focusing on characteristics that implicate the Confrontation Clause's concerns about dignity and reliability. Third, it considers how these underlying characteristics dictate which proposed confrontation rights would be most meaningful.

A machine-specific approach like the one proposed here serves three important functions. First, by considering how machine confrontation would work, it rebuts the judicial argument that such confrontation is impossible or impractical. Second, it offers practical guidance for defendants and cash-strapped attorneys who must face these machines. Third,

* J.D. Candidate, Stanford Law School, 2024. I would like to thank Professor David Alan Sklansky for supervising this endeavor, for inspiring me to write on this topic, and for all his help along the way. I am also incredibly thankful for the feedback, time, and encouragement I received from Yali Corea-Levy, Christopher Garcia, Todd Venook, and Professor David Freeman Engstrom. I also owe a special thanks to Professor Andrea Roth, both for taking the time to discuss this piece with me and for paving the way with her own scholarship in this area. Finally, thank you to Lauryn Bennett and the rest of the *Stanford Law Review* editing team for their excellent comments and feedback.

it sets the stage for other interventions—if judges continue to reject the Confrontation Clause as the vehicle to respond to the problems of machine-generated evidence, legislatures and rulemaking bodies will need guidance on how to structure alternatives.

Table of Contents

Introduction

In a 2019 Pennsylvania assault trial, prosecutors introduced a machine-generated report from the gunshot detection program ShotSpotter.[1] They then relied on that report to show that the defendant, Angelo Weeden, had fired a gun twice in a particular location at a particular time.[2] Weeden appealed his conviction, arguing that he did not have an opportunity to meaningfully confront this machine-generated report, in violation of his rights under the Confrontation Clause.[3] But the Superior Court of Pennsylvania rejected this argument, stating that "it is not possible to cross-examine the declarant of the ShotSpotter report because it was automatically generated by a computer system," and "no one individual can be considered its author."[4] As it turns out, ShotSpotter frequently misclassifies sounds as gunshots, with one Massachusetts police station reporting that the program worked less than 50% of the time.[5] ShotSpotter employees also routinely modify results at law enforcement's request.[6] The court could have granted Weeden access to ShotSpotter's error rates in discovery, giving him the chance to expose these issues. Or the court could have allowed Weeden to call ShotSpotter's programmers to the stand. Instead, the court dismissed machine confrontation as impossible.

Weeden's experience is no anomaly. A New Mexico court recently denied a Confrontation Clause challenge against the statistical output from DNA-testing tool PopStats on the grounds that a mathematical computer accusation cannot be testimonial.[7] PopStats—this purportedly neutral mathematical process—was in the news in 2015 when the FBI discovered that the software had contained an error since 1999 that implicated thousands of cases.[8] Courts

1. Commonwealth v. Weeden, 253 A.3d 329, 332-33 (Pa. Super. Ct. 2021), *aff'd*, 304 A.3d 333 (Pa. 2023).
2. *See id.* at 334.
3. *Id.* at 335-36.
4. *Id.* at 336. It is important to note that the court also held that the report's primary purpose was not use in a later criminal prosecution, which would have been an independent reason to reject Weeden's claim. *Id.* But the court presented the impracticality of machine confrontation as if it were an independent justification, and it is easy to imagine a ShotSpotter report generated after arrest for the purpose of use at trial. *See id.*
5. Garance Burke, Martha Mendoza, Juliet Linderman & Michael Tarm, *How AI-Powered Tech Landed Man in Jail with Scant Evidence*, AP NEWS (Mar. 5, 2022, 1:23 PM EDT), https://perma.cc/YKX2-XS83.
6. *Id.*
7. State v. Espinoza, 525 P.3d 429, 439-40 (N.M. Ct. App. 2022).
8. *See* Spencer S. Hsu, *FBI Notifies Crime Labs of Errors Used in DNA Match Calculations Since 1999*, WASH. POST (May 29, 2015, 10:03 PM EDT), https://perma.cc/4HCC-SVT6.

also routinely deny source code[9] disclosure requests for other machines, despite known cases of software errors in these technologies.[10] And Confrontation Clause challenges are rejected for a wide range of machine-generated evidence.[11] To be sure, allowing confrontation would not necessarily uncover errors in every case. But immunizing machines as neutral tools that cannot be confronted comes at a cost when they are so often incorrect or misleading.

Indeed, the legal and practical questions of machine confrontation have become particularly urgent. Machine-generated evidence is now ubiquitous in criminal trials, and more sophisticated forms of inculpatory evidence are on the way. For decades, courts have grappled with how to handle evidence such as DNA test results and breathalyzer readouts.[12] Despite their long-standing use, these technologies remain a constant source of controversy. Time and time again, further investigation of courtroom technologies that have helped put thousands behind bars has revealed inaccuracies and mistakes.[13] And these are the easy cases. Facial recognition, automated lip reading, automobile drowsiness detection, and drone surveillance are fast approaching, if they are not already here.[14]

Despite this problematic technological expansion, our evidentiary toolkit is ill-equipped to handle the task ahead. Courts have frequently held that the Confrontation Clause does not apply to machine-generated evidence, except in narrow cases where the evidence contains testimonial assertions from human

9. Source code refers to the underlying set of textual instructions written by programmers to control computer programs.

10. *See, e.g.*, People v. Wakefield, 195 N.E.3d 19, 30-31 (N.Y. 2022) (rejecting the argument that the Confrontation Clause requires source code disclosure); Rebecca Wexler, *Life, Liberty, and Trade Secrets: Intellectual Property in the Criminal Justice System*, 70 STAN. L. REV. 1343, 1346 (2018) (collecting cases); David Murray, *Queensland Authorities Confirm 'Miscode' Affects DNA Evidence in Criminal Cases*, COURIER MAIL (Mar. 20, 2015, 10:00 PM), https://perma.cc/PGC4-HRCK (discussing how a software error influenced DNA matches in sixty Australian cases).

11. *See infra* Parts I.B, II.A.

12. *See* Patrick W. Nutter, Comment, *Machine Learning Evidence: Admissibility and Weight*, 21 U. PA. J. CONST. L. 919, 925 (2019).

13. Stacy Cowley & Jessica Silver-Greenberg, *These Machines Can Put You in Jail. Don't Trust Them.*, N.Y. TIMES (Nov. 3, 2019), https://perma.cc/MU27-YNHG (discussing 30,000 New Jersey and Massachusetts breath tests that were thrown out in a single year because of human error and lax oversight); Matt Chapman, Natalie Frazier & The TRiiBE, *False Alarms*, CHI. READER (June 9, 2022), https://perma.cc/ST8T-3J4J (discussing a University of Chicago report finding that 80% of a particular vendor's ankle-monitor alerts were "false positives"); Christian Chessman, Note, *A "Source" of Error: Computer Code, Criminal Defendants, and the Constitution*, 105 CALIF. L. REV. 179, 196-97 (2017) (discussing source code errors detected in the Alcotest 7110 and the Intoxilyzer 5000EN breathalyzer devices).

14. *See infra* Part II.A.

operators.[15] The Supreme Court has obliquely addressed the issue in several of its recent cases, but its holdings suggest only a limited right to cross-examine a human operator that has made an interpretive assertion about raw machine data.[16] Lower courts have spoken more clearly, rejecting the notion that machines are witnesses against a defendant within the meaning of the Confrontation Clause and arguing that machines cannot practically be confronted.[17] Instead, these courts say, defendants should rely on the rules of authentication.[18] In so doing, these courts overlook the low threshold of authentication and the unique accuracy, reliability, and transparency problems presented by machines.[19] Courts make similar arguments against expanding hearsay and due process rights to encompass machine-generated evidence.[20]

Several scholars have argued that courts should read the Confrontation Clause more broadly to consider machine accusers as "witnesses" that trigger "confrontation" rights beyond traditional live testimony.[21] The Confrontation Clause was created to guard against the common law practice of introducing "unconfrontable but impressive-looking" ex parte affidavits.[22] Directly confronting a witness gives defendants the chance to correct mistakes hidden beneath these formal trappings and restore their sense of dignity and fairness. Cross-examination is the specific confrontation right that has persisted in the American legal system, but it is not the only form that would have been considered at common law.[23] Nor should it be the only form now: A host of confrontation rights have been proposed, including source code disclosure, live testimony from the human operator or programmer, broadened discovery

15. *See* Brian Sites, *The Future of the Confrontation Clause: Semiautonomous and Autonomous Machine Witnesses*, 22 VAND. J. ENT. & TECH. L. 547, 562 (2020).

16. *See* Melendez-Diaz v. Massachusetts, 557 U.S. 305, 311 (2009) (holding that an analyst who certifies a forensic drug test must be called to the stand unless unavailable and previously cross-examined); Bullcoming v. New Mexico, 564 U.S. 647, 673-74 (2011) (Sotomayor, J., concurring in part) (leaving open the question of whether a purely machine-generated report requires confrontation).

17. *See infra* Part I.B.

18. *See, e.g.*, United States v. Washington, 498 F.3d 225, 231 (4th Cir. 2007) ("Any concerns about the reliability of such machine-generated information is addressed through the process of authentication not by hearsay or Confrontation Clause analysis.").

19. *See infra* notes 72, 81 and accompanying text.

20. *See generally* Andrea Roth, *Machine Testimony*, 126 YALE L.J. 1972 (2017) (discussing the application of hearsay and other evidentiary rights to machine testimony in detail).

21. *See infra* Part I.C.

22. Roth, *supra* note 20, at 2041-42.

23. David Alan Sklansky, *Hearsay's Last Hurrah*, 2009 SUP. CT. REV. 1, 66; *see* Andrea Roth, *What Machines Can Teach Us About "Confrontation,"* 60 DUQ. L. REV. 210, 217 (2022) ("At common law, cross-examination was neither guaranteed nor deemed sufficient to satisfy the right of confrontation.").

rights over a machine's properties and processes, and the opportunity to interact directly with a machine.[24]

This Note presumes that such machine confrontation rights should exist and focuses on defining meaningful machine confrontation. It argues that meaningful machine confrontation must consider the underlying technology at issue. Existing calls for machine confrontation propose blanket solutions, with limited consideration of the practical concerns judges invoke, the realities defense attorneys face, or the types of technology involved.[25] Potentially testimonial machine accusations are often treated as a single subgroup, particularly when they involve sophisticated algorithms.[26] These schemes accordingly fail to account for the wildly different inputs, outputs, and logic inherent to the range of technologies before the courts. But these specifics matter. Source code disclosure, for example, may have been impractical for interrogating ShotSpotter in Weeden's case because the program uses machine learning and employs decisionmaking that is learned rather than codified.[27] Pretrial disclosure of statistical properties like error rates probably would have been more appropriate. In contrast, source code disclosure may have been precisely what was needed to find material errors in DNA-matching software like PopStats.[28]

This Note offers three contributions to this discussion. First, it provides a comprehensive overview of the technologies entering criminal trials, including those sophisticated tools with the potential to transform the courtroom. Second, it proposes a taxonomy of technological characteristics underlying these tools, focusing on the characteristics that implicate the Confrontation Clause's concerns about dignity and reliability. Finally, it considers how these underlying characteristics dictate which confrontation rights would be most meaningful for defendants.

A machine-specific approach like the one proposed here serves several important functions. By carefully considering how machine confrontation would work, a machine-specific approach rebuts the judicial narrative that vindicating this right is impossible or impractical. This does not resolve all arguments against applying the Confrontation Clause to machine-generated evidence, but it addresses a significant barrier. Analyzing the underlying technologies also offers practical guidance for defendants and cash-strapped

24. *See infra* Part I.D.

25. *See id.*

26. For example, almost all the technologies discussed in Part II.A below would qualify as "Litigation-Related Gadgetry and Software" or "Complex Algorithms" under Andrea Roth's classification. *See* Roth, *supra* note 20, at 2013-22.

27. *See infra* Part III.C; *infra* notes 112-14 and accompanying text.

28. *See supra* notes 7-10 and accompanying text.

attorneys who must face these machines. Regardless of whether this taxonomy has a constitutional dimension, attorneys must be able to understand increasingly technical evidence to adequately address it in court. Finally, practical machine specificity sets the stage for other interventions. If judges continue to reject the Confrontation Clause as the vehicle to respond to the ills of machine-generated evidence, legislatures and rulemaking bodies will need guidance on how to structure alternatives.

This Note proceeds in four Parts. Part I details the narrow and formalistic way courts have treated machine-generated evidence and the Confrontation Clause. It also considers the counterarguments to this treatment and alternative approaches to machine confrontation beyond traditional cross-examination. Part II considers the modern landscape of machine-generated evidence and proposes a set of design characteristics to help assess each technology. Part III discusses how each proposed form of confrontation is influenced by the presence or absence of these characteristics. Finally, Part IV returns to the judicial critiques of machine confrontation and considers how a machine-specific approach answers these concerns. It argues that these practical considerations reveal both that machine confrontation is possible and that some forms of confrontation may be constitutionally inadequate in the face of certain design characteristics. To be sure, there are vexing questions of constitutional line drawing in a machine-specific approach, but they are no different from the kinds of difficult administrability questions faced in other areas of law. The Note concludes by arguing that legislatures can and should intervene using this guidance if courts remain unwilling to rethink their approach to the Confrontation Clause.

I. Current Treatment of Machine-Generated Evidence and the Confrontation Clause

The Confrontation Clause of the Sixth Amendment requires that "[i]n all criminal prosecutions, the accused shall enjoy the right . . . to be confronted with the witnesses against him."[29] Two important questions must be asked to understand how the Clause applies to a piece of machine-generated evidence. Can a machine be a witness against a criminal defendant? And if so, what form of confrontation follows?

This Part presents the argument for treating machines as witnesses that trigger constitutional confrontation protections. Part I.A discusses the evolution of Confrontation Clause doctrine in *Melendez-Diaz v. Massachusetts*[30] and its progeny, highlighting the Supreme Court's reticence towards machine

29. U.S. CONST. amend. VI.

30. 557 U.S. 305 (2009).

witnesses and its dogged commitment to live, human testimony as the appropriate form of confrontation. These cases are the closest the Court has come to addressing purely machine-generated evidence. Part I.B discusses the lower courts, which have more clearly held that machines are not witnesses and cannot be confronted in any practical sense. Several scholars have countered that the Confrontation Clause can or should more comprehensively encompass machine-generated evidence, particularly given the Clause's historical purpose. This argument is briefly sketched in Part I.C. Finally, Part I.D summarizes what types of machine confrontation may be available to criminal defendants if courts honored such a right.

A. *Melendez-Diaz* and its Progeny

For decades, the Supreme Court interpreted the Confrontation Clause as granting criminal defendants the right to cross-examine any hearsay declarant whose introduced statement lacked "adequate 'indicia of reliability.'"[31] But the Court rejected this reliability approach in *Crawford v. Washington*, holding that the central inquiry in deciding whether the opportunity for confrontation is required is whether a statement is "testimonial."[32] A statement is testimonial when the "primary purpose" for its creation was to use it in later criminal proceedings.[33] If a statement is testimonial, the Constitution demands the opportunity for confrontation of the statement's author unless they are unavailable *and* the defendant had a prior opportunity to cross-examine them.[34] *Crawford* emphasizes that confrontation "is a procedural rather than a substantive guarantee" commanding "testing in the crucible of cross-examination."[35]

The Court soon had occasion to apply this test to modern forensic evidence. In *Melendez-Diaz v. Massachusetts*, the Court considered whether a defendant was entitled to confront the analysts who had signed affidavits reporting that the white powder in the defendant's possession was cocaine.[36] The Court determined that this was a "straightforward application of [its] holding in *Crawford*" and held that the affidavits were testimonial statements requiring confrontation.[37] The dissent chafed at this application, decrying the

31. Ohio v. Roberts, 448 U.S. 56, 66 (1980), *overruled by* Crawford v. Washington, 541 U.S. 36 (2004).

32. 541 U.S. at 68-69.

33. Davis v. Washington, 547 U.S. 813, 822 (2006); Ohio v. Clark, 576 U.S. 237, 244-45 (2015).

34. *See Crawford*, 541 U.S. at 59.

35. *Id.* at 61.

36. 557 U.S. 305, 307-08, 311 (2009).

37. *Id.* at 311-12.

administrability challenges of deciding which of many analysts must be cross-examined.[38] This tension over whose conduct produces a testimonial statement in scientific, multi-person evidence generation previews a central challenge for deciding who should testify about machine accusers created by large teams of programmers.[39]

The boundaries of forensic statements were considered again in *Bullcoming v. New Mexico*.[40] At issue was whether the Confrontation Clause could be satisfied by calling a surrogate analyst to testify in lieu of the analyst who had originally signed the affidavit certifying the defendant's blood-alcohol concentration.[41] The Court thought that this affidavit resembled the one in *Melendez-Diaz* "[i]n all material respects."[42] Because the affidavit's author was the witness against the defendant, the author was the one required to testify.[43] As in *Crawford*, the only exception would be if the analyst were legally unavailable and had previously been cross-examined by the defendant, neither of which was true in *Bullcoming*.[44]

Bullcoming most clearly raises the specter of purely machine-generated evidence. The Supreme Court of New Mexico had held that the Confrontation Clause was not violated because the original analyst merely "transcribed the results generated by the gas chromatograph machine."[45] But the Supreme Court rejected this perspective, noting that the analyst had made several independent assertions in the report, such as certifying that he had received the blood sample with the seal still intact.[46] In this way, the Court was able to avoid answering how true machine transcription would be handled. But Justice Sotomayor hinted in concurrence that a purely machine-generated report would be treated differently, noting that "this is not a case in which the State introduced only machine-generated results, such as a printout from a gas chromatograph."[47]

Unfortunately, when the Court soon returned to the issue of forensic evidence in *Williams v. Illinois*, it did little to clarify its approach.[48] In a

38. *Id.* at 332 (Kennedy, J., dissenting) ("Consider how many people play a role in a routine test for the presence of illegal drugs. . . . It is not at all evident which of these [people] is the analyst to be confronted under the rule the Court announces today.").

39. *See infra* notes 185-87 and accompanying text.

40. 564 U.S. 647 (2011).

41. *Id.* at 657-59. The original analyst was on unpaid leave. *Id.* at 659.

42. *Id.* at 664.

43. *Id.* at 659, 663.

44. *See id.* at 659. Here, the prosecution did not assert that the original analyst was unavailable. *Id.*

45. *Id.* (alteration in original) (quoting State v. Bullcoming, 226 P.3d 1, 8 (N.M. 2010)).

46. *Id.* at 659-60.

47. *Id.* at 673 (Sotomayor, J., concurring in part).

48. 567 U.S. 50 (2012).

fractured opinion, a plurality held that the defendant's rights were not violated when an expert testified about a DNA report authored by a non-testifying analyst.[49] Further, these four members of the Court noted that the DNA report "was not prepared for the primary purpose of accusing a *targeted* individual,"[50] suggesting that whether an accusation specifically targets someone may affect what kinds of machine-generated evidence can or should be subject to confrontation.[51] Notwithstanding this observation, *Williams* does little to resolve the treatment of purely machine-generated evidence.[52]

Melendez-Diaz, *Bullcoming*, and *Williams* offer two important insights. First, the Supreme Court has not established clear guidance on how to decide if machine-generated evidence includes testimonial statements, particularly when the outputs are exclusively authored by a machine. The level of involvement humans have in authoring the final output clearly has some bearing, though it may not be dispositive. Nor has the Court confronted hard epistemic questions like whether machines *ever* exclusively author statements. Second, the Court's rigid, formalistic approach to forensics suggests at most a limited, traditional cross-examination right for machine-generated evidence. But the *Melendez-Diaz* dissent's concerns about knowing who to call to testify will only be more pronounced for machines. These cases thus pose a challenge to reimagining machine confrontation rights, but they do not foreclose it.

B. Machine-Generated Evidence in Lower Courts

Lower courts, on the other hand, have had many occasions to consider purely machine-generated evidence. With few exceptions, courts have held that machine-generated evidence does not trigger the Confrontation Clause.[53] Courts have justified this position on three grounds: (1) machines are not "witnesses" within the meaning of the Sixth Amendment; (2) reliability concerns are properly addressed through evidentiary tools like authentication; and (3) machines cannot be cross-examined.

First, courts frequently argue that machines are not witnesses within the meaning of the Confrontation Clause.[54] Early cases focused on the relationship between the Confrontation Clause and hearsay, arguing either that machines

49. *Id.* at 56-58, 62 (plurality opinion).

50. *Id.* at 84 (emphasis added).

51. *See* Brian Sites, *Rise of the Machines: Machine-Generated Data and the Confrontation Clause*, 16 COLUM. SCI. & TECH. L. REV. 36, 74-76 (2014).

52. *See id.* at 50 (noting that *Williams* "does not shed significant light on the underlying question of when machine-generated results trigger the Confrontation Clause").

53. *See id.* at 51.

54. For extensive treatment of this issue, see *id.* at 51-57.

are not "declarants" or that they cannot make "statements."[55] Although *Crawford* largely decoupled hearsay from the Confrontation Clause, courts have continued to adhere to these earlier decisions on the issue of machine-generated data.[56] Other cases assert that, "[i]n light of the constitutional text and the historical focus," the Clause only concerns human witnesses.[57]

More recently, courts have routinely held that machine-generated evidence is not testimonial and therefore does not require confrontation under *Crawford*.[58] Those courts frequently cite Justice Sotomayor's concurrence in *Bullcoming* to differentiate from the testimonial, human-authored affidavits in that case.[59] Like the majorities in *Melendez-Diaz* and its progeny, then, these lower courts presume that it is the *operator's* statement that transforms inculpatory machine-generated evidence into testimony.[60] Besides the occasional reference to historical understanding, few of these decisions address why machine accusations are categorically different from human accusations and why operator involvement is so crucial. Of course, courts could acknowledge that machines can generate testimonial statements without a human operator and still find that a particular piece of evidence lacks the requisite primary

55. *E.g.*, United States v. Washington, 498 F.3d 225, 231 (4th Cir. 2007) (noting that "the raw data generated by the machines" do not constitute "statements," that machines are not "declarants," and that "no out-of-court statement implicating the Confrontation Clause was admitted into evidence"); United States v. Moon, 512 F.3d 359, 362 (7th Cir. 2008) (noting that "data are not 'statements' in any useful sense," nor "is a machine a 'witness against' anyone"); United States v. Lamons, 532 F.3d 1251, 1263 (11th Cir. 2008) (discussing Federal Rule of Evidence 801(a), which defines a hearsay statement as involving a person).

56. *See, e.g.*, United States v. Hayes, 612 F. App'x 673, 675 (4th Cir. 2015) (per curiam) (citing *Washington* favorably); *see also* Sites, *supra* note 15, at 567-68 (discussing the modern treatment of *Washington*).

57. *Lamons*, 532 F.3d at 1263; *see also* State v. Ziegler, 855 N.W.2d 551, 556 (Minn. Ct. App. 2014) (noting that "the constitutional text and the historical focus of the Confrontation Clause . . . clearly establish that the Confrontation Clause is concerned with human witnesses").

58. *E.g.*, United States v. Hill, 63 F.4th 335, 357, 359 (5th Cir. 2023).

59. *E.g.*, United States v. Summers, 666 F.3d 192, 200, 203 (4th Cir. 2011) (citing Bullcoming v. New Mexico, 564 U.S. 647, 672 (2011) (Sotomayor, J., concurring in part)).

60. *See* United States v. Arce, 49 F.4th 382, 392 (4th Cir. 2022) ("But in general, when 'machines generate[] data . . . through a common scientific and technological process,' the *operators of those machines* do not make a 'statement' under the Confrontation Clause when reporting the data." (alterations in original) (emphasis added) (quoting *Washington*, 498 F.3d at 230)); *Hill*, 63 F.4th at 359 ("Key differences exist between test reports generated by a person's analysis and test reports which are the result of machine analysis.").

purpose of using the statements at a later criminal proceeding.[61] But most courts speak broadly about machines and raw data as a class.

Second, courts frequently assert that credibility concerns around machine-generated evidence are properly addressed through other evidentiary tools like authentication.[62] In doing so, courts presume that authentication can screen out patently unreliable technologies and that any additional balancing of accuracy and reliability should be left to the jury. Implicit is the notion that authentication is superior to confrontation, though these cases do not always explain why.

Third, courts assume that machines cannot be confronted. One formulation of this argument equates confrontation to a cross-examination right requiring live question-and-answer before a jury.[63] Live testimony has long been the preferred mechanism for the American adversarial process, and the *Melendez-Diaz* sequence indicates a preference for cross-examination.[64] Because machines cannot be cross-examined in this way, such an understanding forecloses the possibility of a machine confrontation right. Even though there is reason to doubt that confrontation was originally understood to exclusively require cross-examination,[65] it is unsurprising that lower courts have seized upon this view given the Supreme Court's sharp focus on the live testimony of forensic analysts.

But courts frequently suggest something different—an argument of impracticality, not of form. The Seventh Circuit in *United States v. Moon* evoked an image of an oven in the courtroom:

---

61. This reasoning led one court to reject a Confrontation Clause challenge against the introduction of ankle-monitoring data because the monitoring map was created primarily for post-release monitoring. *See* State v. Gardner, 769 S.E.2d 196, 199 (N.C. Ct. App. 2014).

62. *See, e.g.*, *Washington*, 498 F.3d at 231 ("Any concerns about the reliability of such machine-generated information is addressed through the process of authentication not by hearsay or Confrontation Clause analysis."); United States v. Lizarraga-Tirado, 789 F.3d 1107, 1110 (9th Cir. 2015) ("A machine might malfunction, produce inconsistent results or have been tampered with. But such concerns are addressed by the rules of authentication . . . .").

63. *See, e.g.*, People v. Wakefield, 195 N.E.3d 19, 31 (N.Y. 2022) (describing confrontation as a physical cross-examination and rejecting application to source code).

64. *See* Edward K. Cheng & G. Alexander Nunn, *Beyond the Witness: Bringing a Process Perspective to Modern Evidence Law*, 97 TEX. L. REV. 1077, 1077 (2019) ("For centuries, the foundation of the Anglo-American trial has been the witness."); Sklansky, *supra* note 23, at 51 (discussing "the particular 'form' of *Crawford*'s formalism: the equation of 'confrontation' with 'cross-examination'").

65. *See* Sklansky, *supra* note 23, at 66 ("There is little reason to suppose, therefore, that the phrase 'confronted with the witnesses against him' would have been understood in 1791 as simply a way of referring to cross-examination and nothing more."); Roth, *supra* note 23, at 217 ("At common law, cross-examination was neither guaranteed nor deemed sufficient to satisfy the right of confrontation.").

> Yet how could one cross-examine a gas chromatograph? Producing spectrographs, ovens, and centrifuges in court would serve no one's interests. That is one reason why Rule 703 provides that the expert's source materials need not be introduced or even admissible in evidence. The vital questions—was the lab work done properly? what do the readings mean?—can be put to the expert on the stand.[66]

Through this imagery, the court insists that machine confrontation is both impractical and unfruitful. The court did not foresee legitimate credibility concerns about the machine working correctly, only about the human operator using it properly. And the other "vital question" concerned rote interpretation of the machine readings, nothing more.

Perhaps this was not the court's fault. After all, the machines implicated in that 2008 decision were probably well-established, well-respected technologies used in laboratories across the country. These mostly physical machines—breathalyzers, spectrographs, centrifuges—might have been thought to perform simple, uncontroversial chemical processing with less room for error.[67] Times have changed, and technology has become anything but simple. But the list of decisions that still follow suit for new technologies is long.[68]

C. The Case for Applying the Confrontation Clause to Machine-Generated Evidence

Scholars have countered that the Confrontation Clause can and should encompass machine-generated evidence more broadly than it has in the cases described above. As a starting point, classifying machines as witnesses within the meaning of the Confrontation Clause comports with the historical basis for the Clause's creation. The Clause's drafters sought to address a long-standing practice in which justices of the peace collected ex parte affidavits for use in lieu of live testimony.[69] This created "unconfrontable but impressive-looking" evidence that swayed juries and denied defendants the dignity of looking their accusers in the eye.[70] Denying confrontation for machine-generated evidence "resembles trial by ex parte affidavit," because this "raw data" can easily "be incomplete or implicitly biased, even if sincere or technically accurate," and "computer[s] can

66. 512 F.3d 359, 362 (7th Cir. 2008).

67. But this notion was wrong, even in 2008. *See supra* note 13 and accompanying text.

68. *See, e.g.*, People v. Lopez, 286 P.3d 469, 478 (Cal. 2012) (noting that "unlike a person, a machine cannot be cross-examined"); United States v. Lamons, 532 F.3d 1251, 1265 (11th Cir. 2008) (citing *Moon* approvingly); Commonwealth v. Weeden, 253 A.3d 329, 336 (Pa. Super. Ct. 2021), *aff'd*, 304 A.3d 333 (Pa. 2023) (noting that "it is not possible to cross-examine the declarant of the ShotSpotter report because it was automatically generated by a computer system"); *see also* State v. Huettl, 305 P.3d 956, 962, 964 (N.M. Ct. App. 2012) (citing *Moon* approvingly).

69. Roth, *supra* note 20, at 2041.

70. *Id.*

package data in a very enticing manner."[71] Indeed, machines pose acute "'black box' dangers" that make them particularly inscrutable.[72] Courts, then, are misguided in summarily concluding that the "historical focus" of the Clause is at odds with classifying machines as witnesses against defendants.[73]

Of course, a "witness" within the Clause's purview must still make a *testimonial* statement, but many machines arguably do so. Some argue that "a machine source does not make a 'solemn declaration' for the 'purpose' of establishing facts," challenging the notion that a machine could ever satisfy *Crawford*.[74] This argument construes "solemn declaration" and "purpose" narrowly. After all, it is not clear that "purpose" requires machine intent.[75] Indeed, the primary purpose inquiry is objective, suggesting that individual intent is irrelevant.[76] Moreover, some machines are certainly used with the primary purpose of establishing facts in trial. If police run blurry surveillance footage through facial recognition *after* arresting a suspect and try to introduce the resulting printout, the primary purpose of that scan was to produce evidence for trial.[77] And as Andrea Roth aptly notes, "[i]f the point of targeting solemnity is to capture what is particularly abusive about the state purposely relying on impressive but unconfronted allegations," it is hard to imagine machines, with their scientific and seemingly neutral trappings, should be categorically excluded.[78]

Even if recent jurisprudence has rejected the application of the Confrontation Clause to machine-generated evidence, there is no reason that this approach cannot or should not be revisited. For one thing, the lower courts' approach remains an open question for the Supreme Court.[79]

---

71. *Id.* at 2043 (alteration in original) (quoting Jerome J. Roberts, *A Practitioner's Primer on Computer-Generated Evidence*, 41 U. CHI. L. REV. 254, 274 (1974)); *see also Lopez*, 286 P.3d at 494 (Liu, J., dissenting) (noting that the allure of technology might "prompt us to remain alert to constitutional concerns, lest we gradually recreate through machines instead of magistrates the civil law mode of ex parte production of evidence").

72. *See* Roth, *supra* note 20, at 1977.

73. *See supra* note 57 and accompanying text.

74. *See* Roth, *supra* note 20, at 2047.

75. *See id.* at 1988-89 (comparing the question of machine intent to the First Amendment context, which recognizes that determining what constitutes "speech" is not a question of intent, but rather a normative look at why we classify "speech" in the first place).

76. Michigan v. Bryant, 562 U.S. 344, 360 (2011).

77. This kind of "post-targeting" for suspect identification has been at issue in the DNA context already, and at least one court has found that this kind of machine-generated report was testimonial. *See* Young v. United States, 63 A.3d 1033, 1048 (D.C. 2013).

78. Roth, *supra* note 20, at 2048.

79. *See id.* at 2046 ("But even some of the current Justices appear to recognize that the application of the Clause to so-called 'raw data generated by a machine' is an open

footnote continued on next page

Moreover, constitutional doctrine in areas like the Fourth Amendment has evolved with technology, and there is no reason that the same cannot happen here, particularly in light of the Court's admonition that "[r]estricting the Confrontation Clause to the precise forms against which it was originally directed is a recipe for its extinction."[80]

Nor are authentication or other evidentiary rules alone able to fully combat this machine form of ex parte affidavits. The bar for authentication is particularly low and can be met for even the most troubling technologies.[81] Of course, the defense can always challenge the weight of the evidence after it has been introduced. But the credibility problems implicated by machines are complex, and they are often too resource-intensive to counter merely through subpoena.[82] Indeed, *Melendez-Diaz* squarely criticized the burden-shifting to the defense that results absent a constitutional guarantee.[83] Finally, constitutionalizing these protections ensures higher standards of review and more uniform application in the states.[84]

Taken together, these arguments present a strong case that a broad range of machine-generated evidence—including some "purely" machine-generated evidence—should trigger the Confrontation Clause. These arguments also respond directly to the judicial suggestion that machines are not witnesses and that other evidentiary rules are adequate to protect defendants' interests. That leaves only one major critique unaddressed: that machines cannot be confronted.

D. Proposed Forms of Machine Confrontation

There are several forms of machine confrontation that courts could require or allow. The most obvious form would conform to the status quo: live testimony from a human operator. Calling the machine's programmers to the stand is also an option. Doing so raises administrability concerns, much like those mentioned by the dissent in *Melendez-Diaz* regarding which analyst to

question with a nonobvious answer . . . ." (quoting Bullcoming v. New Mexico, 564 U.S. 647, 674 (2011) (Sotomayor, J., concurring in part))).

80. Davis v. Washington, 547 U.S. 813, 831 n.5 (2006).

81. *See* John P. LaMonaca, Comment, *A Break from Reality: Modernizing Authentication Standards for Digital Video Evidence in the Era of Deepfakes*, 69 AM. U. L. REV. 1945, 1976-77 (2020) (discussing the low bar of authentication in relation to video manipulation technology).

82. *See* Sites, *supra* note 15, at 571.

83. *See* Melendez-Diaz v. Massachusetts, 557 U.S. 305, 324 (2009) ("Unlike the Confrontation Clause, [subpoenas] are of no use to the defendant when the witness is unavailable or simply refuses to appear.").

84. *See* Sites, *supra* note 15, at 574.

call to the stand.[85] Still, some scholars have argued that "[a]ll output from a computer program constitutes a statement that is authored (at least in part) by the computer scientist," and a confrontation right should properly attach.[86] Robust forms of this right could require a programmer to testify before a scientific commission every time the software changes.[87]

A growing number of scholars, attorneys, and activists have called for broader access to machine source code. Some have framed this call in the context of the Confrontation Clause.[88] Others have argued that source code access should be a precondition for admitting computer evidence.[89] While judges have broadly rejected source code as either irrelevant or protected under trade secret privilege, some have occasionally granted disclosure.[90] Where permitted, disclosure has proven useful: Source code analysis of courtroom technology has led to the discovery of software errors in several cases.[91]

Courts could also expand discovery to accommodate pretrial confrontation. This could work in several ways: through pretrial interrogatories "cross-examining" the machine on statistics like error rates;[92] disclosure of prior machine statements;[93] or disclosure of machine documentation, calibration information, and business records.[94] The defendant

85. *Melendez-Diaz*, 557 U.S. at 332 (Kennedy, J., dissenting) ("It is not at all evident which of these four persons is the analyst to be confronted under the rule the Court announces today.").

86. Chessman, *supra* note 13, at 220-21. *But see* Roth, *supra* note 20, at 1988 (noting the challenges of understanding programmer authorship and arguing that "any ruling allowing the programmer to testify should not be based on the premise that the programmer is the true declarant").

87. *See* Roth, *supra* note 20, at 2050. This kind of structural solution will largely remain undiscussed in this Note.

88. *See id.*

89. *See, e.g.*, Wexler, *supra* note 10 (discussing source code disclosure and its value at length); Chessman, *supra* note 13, at 221 ("Requiring disclosure of source code as a condition of admissibility ensures that prosecutors cannot take advantage of the benefits of computer programs without the programs being subjected to the rigors of adversarial testing."); Edward J. Imwinkelried, *Computer Source Code: A Source of the Growing Controversy over the Reliability of Automated Forensic Techniques*, 66 DEPAUL L. REV. 97, 101 (2016) (discussing source code disclosure in the context of DNA analysis and arguing for a limited defense right of access).

90. *See* Imwinkelried, *supra* note 89, at 100-01 & nn.21-22; Wexler, *supra* note 10, at 1358-62.

91. *See, e.g.*, Imwinkelried, *supra* note 89, at 119 & n.168 (citing a report submitted in *State v. Chun*, 943 A.2d 114, 132-33 (N.J. 2008)); Roth, *supra* note 20, at 1999 n.134.

92. *See* Roth, *supra* note 20, at 2050.

93. *Id.*

94. *See* Jennifer Mnookin & David Kaye, *Confronting Science: Expert Evidence and the Confrontation Clause*, 2012 SUP. CT. REV. 99, 106, 125 n.85, 156 (addressing machine documentation, specifically); Cheng & Nunn, *supra* note 64, at 1106 (discussing these forms of enhanced discovery outside the context of machine-generated evidence).

could also receive access to the actual machine to tinker with, either in pretrial discovery or in front of the jury.[95]

Deciding which form of machine confrontation to use depends partially on how willing courts are to decouple confrontation from in-court cross-examination. While courts have largely hewn towards a traditional approach, there is reason to doubt—as a constitutional matter—that confrontation can only be satisfied by cross-examination.[96] Besides, "there is little reason to narrowly construe" the Confrontation Clause "as guaranteeing only the courtroom safeguards of the oath, physical confrontation, and cross-examination" when dealing with such novel ground.[97]

Instead, the focus should be on deciding which form of confrontation will achieve the Clause's original aims: promoting reliability, fairness, and transparency.[98] Scholars have focused thus far on whether a given method is *generally* an effective approach for machine-generated evidence. Few, if any, have considered how the *type* of technology—which can range from a simple electrochemical reaction in a breathalyzer to a complicated, neural network in a facial recognition algorithm—dictates when a method effectively achieves those aims. Just like different witnesses merit different approaches in cross-examination, so too do different machines. Particularly where overworked and underpaid public defenders must make difficult choices about how to allocate a finite pool of resources towards unfamiliar and complex technologies, these differences matter for defendants confronting potentially inaccurate or unreliable machines.[99]

II. Charting the Landscape of Machine-Generated Evidence

Technology is now used at nearly every phase of criminal proceedings.[100] The range of technologies is expansive, encompassing everything from the

95. *See* Cheng & Nunn, *supra* note 64, at 1107; Jennifer L. Mnookin, *Repeat Play Evidence: Jack Weinstein, "Pedagogical Devices," Technology, and Evidence*, 64 DEPAUL L. REV. 571, 588 (2015) ("The simulation should be provided in a form that lets the party modify the assumptions, and perhaps even change aspects of the underlying computer code . . . and see how these changed assumptions change the result, if at all.").

96. *See supra* note 65 and accompanying text.

97. Roth, *supra* note 20, at 2048.

98. *See infra* notes 169-74 and accompanying text.

99. *See* Erin Murphy, *The New Forensics: Criminal Justice, False Certainty, and the Second Generation of Scientific Evidence*, 95 CALIF. L. REV. 721, 763 (2007) (noting that, in some cases, "the defense attorney with limited resources would be remiss, both practically and ethically, in wasting precious time and effort carefully opposing the admission of the scientific evidence").

100. Wexler, *supra* note 10, at 1346 (noting that "[a]t every stage-policing and investigations, pretrial incarceration, assessing evidence of guilt at trial, sentencing, and parole-

footnote continued on next page

mundane (i.e., breathalyzers) to the futuristic (i.e., artificial intelligence). This pervasiveness makes classifying these technologies particularly urgent, but the breadth makes such a task exceedingly difficult. Part II.A begins by surveying the broad range of machine-generated evidence, particularly those technologies that are only just entering the courtroom. These technologies are frequently treated cursorily in scholarly introductions, but understanding the spectrum is key to extracting a workable taxonomy. Part II.B then considers machine design and proposes a set of technological characteristics that bear heavily on machine accuracy, reliability, and transparency. It argues that these characteristics should be a north star in understanding which method of confrontation should be employed for a given machine.

A. Modern Machine-Generated Evidence

Courts are not just seeing ovens and spectrographs anymore. They are already seeing a wide range of sophisticated artificial intelligence, complex chemical calculations, and multi-factor automobile awareness systems that were once the stuff of science fiction. These tools do not just have a wide range of purposes—they have a wide range of *underlying technological designs*. Design choices have received some discussion,[101] but they have largely been glossed over or ignored outright. This Subpart takes a closer look at this range of designs. Not every technology listed here would always warrant confrontation, but this discussion demonstrates the wide array of available technologies and illustrates why a one-size-fits-all confrontation solution is unlikely to be successful.

Forensic Evidence. Perhaps the most familiar form of machine-generated evidence introduced at trial is forensic evidence like DNA analysis,[102] drug testing results,[103] and blood alcohol testing results.[104] Such evidence has occupied center stage in discussions of the primary purpose test[105] and source code disclosure.[106] However, these types of evidence are governed by wildly

machine learning systems and other software programs increasingly guide criminal justice outcomes"); Sites, *supra* note 15, at 549-50 (describing the various contexts in which machines can be witnesses in criminal proceedings).

101. *See, e.g.*, Roth, *supra* note 20, at 1990-2000.

102. *See, e.g.*, United States v. Summers, 666 F.3d 192, 195-96 (4th Cir. 2011).

103. *See, e.g.*, Melendez-Diaz v. Massachusetts, 557 U.S. 305, 307 (2009); United States v. Blazier, 69 M.J. 218, 221 (C.A.A.F. 2010).

104. *See, e.g.*, People v. Lopez, 286 P.3d 469, 471 (Cal. 2012); People v. Dinardo, 801 N.W.2d 73, 75 (Mich. Ct. App. 2010).

105. *See supra* Part I.A.

106. *See, e.g.*, David Liebow, Note, *DWI Source Code Motions After* Underdahl, 11 MINN. J.L. SCI. & TECH. 853, 853 (2010) (discussing source code requests for breathalyzers); Wexler, *supra* note 10, at 1392 (discussing the history of DNA source code disclosure requests).

different technologies. Breathalyzers, for example, detect ethanol concentration by shining an infrared light through a breath sample and reporting light absorption rates.[107] In contrast, DNA analysis employs a technique called probabilistic genotyping, which creates a raw DNA profile from a sample using a series of chemical reactions.[108] A proprietary software system like STRmix or TrueAllele then calculates the likelihood that a person's DNA is in that sample profile.[109] Chemical reactions are central to each of these technologies, but they are otherwise quite distinct. Nevertheless, as was the case for PopStats, each of these types of forensic evidence has come under fire for miscalibration, misuse, and inaccuracy.[110]

Location-Based Technology. Ankle monitors, handheld GPS devices, and Stingrays (secretive tools that mimic cellphone towers to receive phone location data) are frequently used to track a defendant's location.[111] The controversial tool ShotSpotter uses acoustic sensors installed on buildings and telephone poles to triangulate sounds that machine learning predicts to be gunfire.[112] This technology incorrectly flags everything from church bells to trash pickup, noises are regularly relabeled by employees at law enforcement's request, and location data is sometimes changed by employees.[113] Despite these inaccuracies, ShotSpotter has been admitted in court, for instance, in Angelo Weeden's case.[114]

Electronic Records. Another major category of machine-generated evidence is electronic records, including auto-generated receipts[115] and computer

107. Cowley & Silver-Greenberg, *supra* note 13.

108. WILLIAM C. THOMPSON, SPECIAL MASTER'S REPORT ON THE SCIENTIFIC FOUNDATIONS OF STRMIX™ 4, 5, 15 (2019).

109. *Id.* at 24-25 (discussing STRmix); *see also* Justin Jouvenal, *A Secret Algorithm Is Transforming DNA Evidence. This Defendant Could Be the First to Scrutinize It.*, WASH. POST (July 13, 2021, 8:00 AM EDT), https://perma.cc/BMC8-ZRAG (discussing another common tool, TrueAllele).

110. *See supra* notes 10, 13 and accompanying text; *see also* Tracey Kaplan, *Crime Lab Uses Wrong Chemical in 2,500 Methamphetamine Tests in Santa Clara County*, MERCURY NEWS (May 5, 2014, 12:34 PM), https://perma.cc/7CLT-TSFN (discussing similar issues in drug tests).

111. *See, e.g.*, People v. Rodriguez, 224 Cal. Rptr. 3d 295, 296-97 (Cal. Ct. App. 2017) (ankle monitor); United States v. Lizarraga-Tirado, 789 F.3d 1107, 1108 (9th Cir. 2015) (handheld GPS device); United States v. Patrick, 842 F.3d 540, 542-43 (7th Cir. 2016) (Stingrays); Adeline Lee & Laura Moraff, *Surreal Stingray Secrecy: Uncovering the FBI's Surveillance Tech Secrecy Agreements*, ACLU (Dec. 15, 2021), https://perma.cc/UX4V-PCB6.

112. *See* Burke et al., *supra* note 5.

113. *Id.*

114. Commonwealth v. Weeden, 253 A.3d 329, 334 (Pa. Super. Ct. 2021); *see also, e.g.*, United States v. Godinez, 7 F.4th 628, 633 (7th Cir. 2021).

115. *See, e.g.*, United States v. Channon, 881 F.3d 806, 809 (10th Cir. 2018).

logs.[116] Hard drive extraction reports are of particular importance: Law enforcement uses the computer program Cellebrite to download information from a phone or computer hard drive and create a browsable report of call records, location data, and browser history.[117] Cellebrite reports, internet access logs, and computer records also play a prominent role in cybercrime allegations.[118]

Imaging and Video. Trials regularly use surveillance videos, body camera footage, animations, red light traffic camera printouts,[119] and even crime scene simulations.[120] Sophisticated altered evidence—"deepfakes"—has even appeared in a criminal trial.[121] But even before deepfakes, imaging and video technology was deceptively complex and liable to make mistakes.[122]

Artificial intelligence plays an important role in modern imaging and video technology. Computer vision and machine learning, two forms of artificial intelligence, are powering sophisticated facial, audio, and spatial recognition. Lip-reading algorithms can identify words with nearly twice the accuracy of humans.[123] License plate scanners employ machine learning and optical character recognition to parse and interpret passing license plates.[124] Evidence suggests that facial recognition is solely used as an investigative tool for now.[125] The same appears true for these other technologies. But the use case

116. *See, e.g.*, United States v. El Gammal, 831 F. App'x 539, 541-42 (2d Cir. 2020) (discussing Facebook metadata).

117. *See* United States v. Arce, 49 F.4th 382, 392-93 (4th Cir. 2022) (describing a basic Cellebrite report).

118. *See, e.g.*, *id.*; United States v. Wilson, 13 F.4th 961, 964 (9th Cir. 2021) (discussing a child pornography case involving a Google-generated report including internet access information and computer records).

119. *See, e.g.*, People v. Goldsmith, 326 P.3d 239, 242 (Cal. 2014).

120. *Forensic Crime Scene Reconstruction, Virtual Reality*, NAT'L CTR. FOR AUDIO & VIDEO FORENSICS, https://perma.cc/24LY-57AV (archived Apr. 11, 2024) (using laser scanning to recreate crime scenes for juries); Curtis E.A. Karnow, *The Opinion of Machines*, 19 COLUM. SCI. & TECH. L. REV. 136, 158 (2017) (discussing airplane crash simulations).

121. *See* Riana Pfefferkorn, "*Deepfakes" in the Courtroom*, 29 B.U. PUB. INT. L.J. 245, 254 n.40, 263 (2020) (discussing the introduction of a doctored piece of evidence in the United Kingdom, though it is not clear how sophisticated this alleged "deepfake" was).

122. *See* Ted Chiang, *ChatGPT Is a Blurry JPEG of the Web*, NEW YORKER (Feb. 9, 2023), https://perma.cc/7MBX-A8ZR (discussing Xerox scans that failed to reproduce numbers correctly because of the compression algorithm used to store them).

123. Jamie Condliffe, *AI Has Beaten Humans at Lip-Reading*, MIT TECH. REV. (Nov. 21, 2016), https://perma.cc/TMN9-5NGX (noting that human volunteers correctly identified only 52.3% of words, while the machine correctly identified 93.4%).

124. *See, e.g.*, Dep't of Homeland Sec., Automated License Plate Reader (ALPR) (2021), https://perma.cc/48TG-G242 (explaining how automated license plate readers work).

125. *See* People v. Reyes, 133 N.Y.S.3d 433, 436 (N.Y. Sup. Ct. 2020) (suggesting that there are no New York cases involving introduction of a facial recognition "match" at trial).

for these technologies in trial is clear: When image or audio quality of, say, a surveillance feed is poor, machine outputs predicting who is shown or what was said could be introduced at trial.[126]

Automobiles. Automobile technology is likely to see increased use at trial as cars collect more data. A car's sensing and diagnostic module, for example, tracks the car and can help identify a suspect's driving behaviors.[127] Some autonomous vehicles have drowsiness detection systems, which watch facial contours, pedal pressure, steering wheel movement, and other signals to guess when the driver is falling asleep.[128] It is easy to imagine this technology being used as evidence of driving under the influence or reckless driving.

Something More? As technology changes, this list will continue to grow. Some technologies—for example, machine learning models that predict whether someone authored a text—do not clearly fit into any of these categories.[129] And drones,[130] brain imaging analysis,[131] and lie detectors[132] push the boundaries of what might be introduced at trial in the future. Keeping potential future technologies in mind will be crucial to maintaining a practical taxonomy for machine-generated evidence.

B. Classifying Technological Characteristics

An intuitive but non-exhaustive way to think about a machine's design is to consider: (1) the inputs; (2) what the machine does with the inputs; (3) how the machine knows what to do with the inputs; (4) whether that knowledge can

126. Joseph Clarke Celentino, Note, *Face-to-Face with Facial Recognition Evidence: Admissibility Under the Post-*Crawford *Confrontation Clause*, 114 MICH. L. REV. 1317, 1320-21, 1325 (2016) (discussing the future use of facial recognition).

127. *See* Nguyen v. State, No. 05-20-00241-CR, 2022 WL 3714494, at *1-2 (Tex. Ct. App. Aug. 29, 2022) (discussing the use of a crash data report generated by a car's black box recorder).

128. *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) (exploring evidentiary approaches to evidence generated by AI-driven systems, including drowsiness detection systems). The proliferation of self-driving technology will only accelerate this trend. *See also* Kirsten Korosec, *Tesla Has Activated Its In-Car Camera to Monitor Drivers Using Autopilot*, TECHCRUNCH (May 27, 2021, 3:56 PM PDT), https://perma.cc/M4WS-M32K (outlining the driver detection system in Teslas).

129. *See* Andrea Roth, *The Use of Algorithms in Criminal Adjudication*, *in* THE CAMBRIDGE HANDBOOK OF THE LAW OF ALGORITHMS 407, 413-14 (Woodrow Barfield ed., 2021).

130. *See generally* GREGORY MCNEAL, BROOKINGS, DRONES AND AERIAL SURVEILLANCE: CONSIDERATIONS FOR LEGISLATURES (2014), https://perma.cc/9PX7-5PBV (anticipating future concerns about the use of drone surveillance).

131. *See* Emily R.D. Murphy & Jesse Rissman, *Evidence of Memory from Brain Data*, 7 J.L. & BIOSCIENCES lsaa078, at 1 (2020), https://perma.cc/4HUS-MP6D.

132. *See, e.g.*, Frye v. United States, 293 F. 1013 (D.C. Cir. 1923).

change over time; (5) the outputs; and (6) how a human uses the system.[133] This Subpart considers these high-level design questions and extracts a set of characteristics that distinguish the underlying technologies producing machine-generated evidence in criminal trials: (1) multi-dimensionality, (2) transient inputs, (3) machine training, (4) codebase complexity, (5) human discretion, and (6) physical decay. While not an exhaustive list, these characteristics have been chosen because they vary wildly among technologies deployed in court. Each of these machine characteristics also implicates the underlying purposes of the Confrontation Clause. For example, large amounts of human discretion and non-diverse machine learning training sets can both involve bias and reduce neutrality.[134] This Subpart will examine these design characteristics, setting the stage for Part III's consideration of how they shape confrontation's efficacy.[135]

1. Multi-dimensionality and transient inputs: What are the machine's inputs?

The type and number of machine inputs have important implications for a machine's sensitivity and transparency. Multi-dimensional systems take in several machine inputs, which influences the number of factors the machine weighs in decisionmaking. For example, an airplane crash simulation takes in radar data, crash-site measurements, and the airplane's "black box"; and its output is a product of each of these inputs.[136] Similarly, automobile drowsiness detection can consider steering wheel angle, lane deviation, posture, facial expression, and heart rate.[137] Systems taking in more information can have more variable outputs, and it can be harder to identify how each input influenced those outputs.

133. Machine outputs and input transformations go largely undiscussed in this Note, though they are certainly worthy of future discussion.

134. *See* Batya Friedman & Helen Nissenbaum, *Bias in Computer Systems*, 14 ACM TRANSACTIONS ON INFO. SYS. 330, 333-36 (1996) (discussing preexisting, technical, and emergent bias, and noting the various parts of technical design where these forms creep in).

135. Some scholars have developed different schemes for classifying machine-generated evidence. *See* Murphy, *supra* note 99, at 726-30 (distinguishing between first- and second-generation forensic science); Sites, *supra* note 51, at 66-91 (distinguishing based on the type of human involvement in the production of the evidence); Roth, *supra* note 20, at 2006-22 (distinguishing between machines that implicate credibility concerns and those that don't, and sub dividing based on historical treatment). While the characteristics that emerge across these subdivisions—particularly Roth's—sometimes overlap with this taxonomy, existing schemes are too broad and encompass too many machine designs to fully answer which form of machine confrontation to adopt.

136. *See* Karnow, *supra* note 120, at 158.

137. *See* Eric A. Taub, *Sleepy Behind the Wheel? Some Cars Can Tell*, N.Y. TIMES (Mar. 16, 2017), https://perma.cc/M667-9UZT.

Multi-dimensionality is not just a product of the number of discrete inputs. Drowsiness detection, for example, is also multi-dimensional in the sense that *multiple features* of an individual input are considered. There, the machine learning model considering the facial expression parses out a subset of "features" from the raw data that it uses for predictive classifications.[138] A "feature" in this sense describes a recurring pattern that the machine uses to generate a classification, though the features are not always as easily understandable as "nose shape" or "distance between eyes." Feature selection is a critical part of machine learning because reducing the number of features that the model considers during prediction trains the model faster and filters out irrelevant data. It also avoids "over-fitting" to too many variables at once and reducing the predictive power of the model to new data that is not quite as consistent.[139] Crucially, feature selection—reducing the multi-dimensionality, in a sense—can also reduce complexity and make it easier to explain the model's output.[140]

Single-input, one-dimensional systems are difficult to find. Simple thermometers are one example, but such devices may not implicate the hardest confrontation questions because they are mostly reliable and understandable. Ultimately, the primary significance of multi-dimensionality is its complexity, so *particularly* multi-dimensional machines—which are the easiest to identify just by learning about how the machine receives inputs—are most likely to need robust confrontation.

In addition to the number of inputs, the type of input is also an important design choice. Inputs that are *transitory* and only exist for a fleeting moment are difficult to replicate during criminal proceedings. Lie detectors and brain scans are great examples, as the single "input" in these cases is probably near impossible to replicate in an organic fashion, even if the machine could store a logged representation of what it had been at the time. This is distinct from a facial recognition algorithm that takes in a static image and thus can receive that same input at any time. Because transitory inputs are so difficult to

138. *See* Sameer Singh, Maneesha Singh & Markos Markou, *Feature Selection for Face Recognition Based on Data Partitioning*, 1 PROC. 16TH INT'L CONF. ON PATTERN RECOGNITION 680, 683 (2002) (discussing approaches to feature selection for facial recognition); Celentino, *supra* note 126, at 1328 (discussing the use of algorithms to map data points of a person's face); David Lehr & Paul Ohm, *Playing with the Data: What Legal Scholars Should Learn About Machine Learning*, 51 U.C. DAVIS L. REV. 653, 677 n.92 (2017) (discussing the complexity of feature extraction and the wide range of approaches that are used).

139. A model trained on too many irrelevant features will be unduly sensitive to noise and other small changes from the original dataset. *See* Renu Khandelwal, *Feature Selection: Identifying the Best Input Features*, TOWARDS DATA SCI. (Oct. 24, 2019), https://perma.cc/DLB6-ZH8M.

140. *Id.*

replicate, their presence likely implicates different confrontation rights than those that are associated with static systems.

2. Machine training: How does the machine know what to do?

Technologies are also distinguishable by how they are trained. Some machines are told what to do. This Note will call these machines "rule-based," though the term lacks some descriptive value. In these systems, humans write code that specifically instructs the machine on its conditional reasoning: For these inputs, compute this output using that formula.

Other machines are taught how to think by example. Machine learning and other forms of artificial intelligence—the systems partially underlying facial recognition, drowsiness detection, automated lip-reading, and ShotSpotter—fall into this category. These models are given training inputs, often in the form of large datasets with pre-labeled classifications, and they identify patterns across those classifications.[141] For example, a machine might teach itself that certain keywords indicate fraudulent behavior. Often, "[n]o one knows why the system selects [something]: once the system is trained, no script can be provided to a human sorter to imitate the system's selection."[142] Machines can be taught to do everything from classification to anomaly detection to transcription.[143]

The landscape of machine learning (let alone artificial intelligence) is extensive, but a brief discussion should suffice here.[144] One common form of machine learning is *supervised learning*, which requires structured, pre-labeled training data.[145] A very simple example dataset for a self-driving car's vision system would be a set of pictures that were hand-labeled by an analyst as cars, cyclists, or trees.[146] Other models rely on *unsupervised learning*, which does not use pre-labeled data and instead lets the machine "cluster" common patterns that it detects in the training data without ever understanding what properties it is sorting.[147] Facial recognition, for example, uses unsupervised learning.[148] Under both forms of learning, these training datasets are collected either by

141. Nutter, *supra* note 12, at 927-28 (describing how machine learning models learn patterns across "thousands or even millions of examples").

142. Karnow, *supra* note 120, at 142.

143. IAN GOODFELLOW, YOSHUA BENGIO & AARON COURVILLE, DEEP LEARNING 98-100 (2016); Nutter, *supra* note 12, at 929.

144. For more discussion on the topic, see Lehr & Ohm, note 138 above, at 655 (attempting to provide a "rich breakdown of the process of machine learning").

145. GOODFELLOW ET AL., *supra* note 143, at 103; Lehr & Ohm, *supra* note 138, at 673.

146. *See* GOODFELLOW ET AL., *supra* note 143, at 105.

147. *See* Lehr & Ohm, *supra* note 138, at 676.

148. *See* Karnow, *supra* note 120, at 143.

merging existing datasets or through a herculean effort to collect a statistically noiseless, representative dataset.[149] Models must also then be run against a test dataset and tweaked based on the calculated performance.[150]

It should be no surprise that machine learning can introduce bias, inaccuracy, and opacity. Trained models do not have explicit commands that fully govern their operation.[151] Accordingly, opacity is an enormous issue, which has spawned a subdiscipline of academic inquiry.[152] For a complex model, it may not *ever* be possible to identify what led it to classify an input in a certain way.[153] Models can be—and are—trained using biased or incomplete datasets.[154] They can detect patterns that humans did not intend to identify,[155] and their classifications are not immune from mistakes.[156] They may also be sensitive to minute changes.[157]

Rule-based systems are not immune to these risks, though it is probably true that transparency is more difficult to achieve in complex machine-trained models. Perhaps the biggest issue with rule-based systems is that they give programmers subjective control over decisions like threshold values, such as how large a DNA sample must be before it is considered.[158] Accordingly, both types of machines implicate different confrontation challenges.

---

149. *See* Lehr & Ohm, *supra* note 138, at 677-81.

150. *See id.* at 698-99.

151. *See* Karnow, *supra* note 120, at 142-46.

152. *See, e.g.*, DAVID FREEMAN ENGSTROM, DANIEL E. HO, CATHERINE M. SHARKEY & MARIANO-FLORENTINO CUÉLLAR, GOVERNMENT BY ALGORITHM: ARTIFICIAL INTELLIGENCE IN FEDERAL ADMINISTRATIVE AGENCIES 75-78 (2020) (discussing algorithmic transparency in the administrative state).

153. Lehr & Ohm, *supra* note 138, at 708-10; ENGSTROM ET AL., *supra* note 152, at 75-77 (discussing deep learning and other techniques that are less interpretable by design).

154. *See, e.g.*, Olga Akselrod, *How Artificial Intelligence Can Deepen Racial and Economic Inequities*, ACLU (July 13, 2021), https://perma.cc/TD3K-HNXJ; Michael L. Rich, *Machine Learning, Automated Suspicion Algorithms, and the Fourth Amendment*, 164 U. PA. L. REV. 871, 885 (2016) (noting that decisions about training data "allow human assumptions about what correlations *should* exist in the data to color the outcome").

155. Nutter, *supra* note 12, at 951 & n.199 (citing a machine learning model used to differentiate Russian and American tanks that inadvertently used image brightness as the determinative feature).

156. *See* Boris Babic, I. Glenn Cohen, Theodoros Evgeniou & Sara Gerke, *When Machine Learning Goes Off the Rails*, HARV. BUS. REV., Jan.-Feb. 2021, at 76.

157. *See* Lehr & Ohm, *supra* note 138, at 704-05 (discussing models with highly disparate outcomes and techniques to combat sensitive features).

158. *See, e.g.*, Lauren Kirchner, *Traces of Crime: How New York's DNA Techniques Became Tainted*, N.Y. TIMES (Sept. 4, 2017), https://perma.cc/LN7Y-32TN ("To reduce potential problems, the lab decided not to amplify samples smaller than 20 picograms, or about three cells' worth of DNA . . . .").

3. Codebase complexity: How are the instructions to the machine expressed?

The complexity of the underlying source code (codebase complexity) is another influential technological characteristic. Codebase complexity is tightly coupled with transparency and accuracy. Complex code is opaque and difficult for even an expert to parse. And computer "bugs" are easier to cause—and harder to find—when software is spread across a large or dense codebase.[159] This is particularly true when new engineers begin working with complex codebases and lack knowledge about the intricacies of the rest of the system.[160] Some code is complex because it has millions of lines of code, some because it involves complicated and novel algorithms, some because it was poorly written, and some because it is split into files and directories that are deeply nested and difficult to parse.[161]

It is tempting to think that certain technologies are inherently more complex than others and thus codebase complexity might be a byproduct of another design characteristic. After all, a multi-input drowsiness detection system is going to require more code than a basic breathalyzer. But some seemingly simple technology can become complicated purely because the program has existed for a long time and has become file-dense over many updates.[162]

This complexity poses a challenge. The characteristics discussed so far are quick heuristics for determining the correct confrontation right *ex ante*, but complexity may not be readily understood until after a judge or defendant has had an expert review the code. But this possibility is not fatal to the usefulness of this characteristic. Some forms of complexity are more readily detected *ex ante*. Lines of code and number of files, for instance, may be good proxies for complexity.

4. Human discretion: How is the machine used?

Unlike the aforementioned characteristics, human use is already a central consideration in existing Confrontation Clause jurisprudence[163] and proposed machine-classification schemes.[164] This Subpart, however, focuses more

159. *See* Chessman, *supra* note 13, at 186-87.

160. *See id.* at 189-90.

161. For a broader discussion about the subtle ways in which errors creep into complex software programs, see *id.* at 186-96.

162. *See id.* at 190-91.

163. *See supra* notes 45-47 and accompanying text (discussing *Bullcoming*'s apparent holding that operator interpretation of the output was key to finding a Confrontation Clause violation).

164. *See supra* note 135.

narrowly on *discretion*—where human operators make choices between a set of similarly reasonable options. To be clear, a system that involves a lot of human control does not necessarily allow much human discretion. A radar gun, for example, depends on human use but permits almost no human discretion. Calibration is one important form of human discretion, as is the ability to choose the inputs to the system. Choice of input can manifest in subtle ways: for instance, intentionally or accidentally placing a thermometer near a cooler air duct to change the "room" that is being measured,[165] or pre-treating inputs to a facial recognition system by modifying image lighting.[166]

5. Physical decay: Can the machine's reasoning change?

All the characteristics thus far have dealt with a relatively fixed machine design, but machines are liable to change their reasoning over time. Some algorithms incorporate randomness by design and not always in obvious ways.[167] Source code updates, model re-training, and database expansion can all change outputs—and error rates—over time.

Machine reasoning also changes over time because of decay, which is often physical. For instance, the microphone installed in a mounted ShotSpotter device will decay over time, and it may eventually pick up distorted sound waves as a result. There are also many forms of non-physical decay, such as software mistakes that creep in when code is not properly maintained.[168]

Decay alone does not describe an underlying design feature, but there are some technological components that are particularly susceptible to decay. Machines that rely on fragile physical sensors or are exposed to significant environmental pressure are more likely to decay. But determining the likelihood of decay is not always an easy task. Attorneys may struggle to quickly determine the durability of relevant sensors. Assessing the likelihood of decay, then, may require reliance on proxies—for instance, regular maintenance requirements might signal that a machine is particularly at risk.

III. A More Meaningful Right to Confront

What these characteristics imply for "meaningful" confrontation depends on how we conceive of the right's goals. One core goal of confrontation is to "minimize inferential error by giving the jury sufficient context to understand

165. *See* Roth, *supra* note 20, at 1993.

166. Celentino, *supra* note 126, at 1327.

167. *See* Joshua A. Kroll et al., *Accountable Algorithms*, 165 U. PA. L. REV. 633, 647-48 (2017).

168. *See supra* Part II.B.3.

the probative value of the evidence."[169] Cross-examination seeks to achieve this goal in at least three ways, by helping the jury understand: (1) what the witness is saying, (2) whether the witness is telling the truth, and (3) whether the witness is correct. During cross-examination, attorneys facilitate these goals by tailoring their questions to the witnesses before them. When a witness is evasive about their decision-making, the attorney can approach the issue indirectly: What were the witness's past practices? Were there factors that did *not* inform their reasoning? Should the jury view their answers as deceptive and unreliable?[170]

In this regard, a machine is not much different from a human witness: Attorneys can likewise tailor inquiries into transparency, reliability, and accuracy to the design characteristics of the machine witness. Machine confrontation, however, introduces challenges not implicated by human confrontation. Unlike a human on the stand, it is harder and more expensive to shift to a different form of confrontation in the machine context. Source code, discovery, and calling a new witness to the stand are all expensive,[171] and if a defense attorney realizes too late that the human operator or source code cannot answer the question, they will need to move to a different form of confrontation. Machine witnesses will also change over time as technology evolves, and familiar tactics for one tool may not work on another. Assessment of the design characteristics delineated in Part II is one way to respond to this issue. Design characteristics reveal when certain forms of confrontation will be effective at demonstrating—or discrediting—machine reliability for the jury. Thus, design characteristics offer attorneys—especially cash-strapped public defenders—a quick heuristic for the kind of "witness" they face.

Confrontation also implicates concerns of dignity and fairness, and meaningful machine confrontation should consider these as well.[172] Part of the value of cross-examination is that it allows a defendant to look their accuser in

169. Roth, *supra* note 23, at 221; *see also* Maryland v. Craig, 497 U.S. 836, 846 (1990); Dutton v. Evans, 400 U.S. 74, 89 (1970) (plurality opinion) ("[T]he mission of the Confrontation Clause is to advance a practical concern for the accuracy of the truth-determining process in criminal trials by assuring that 'the trier of fact [has] a satisfactory basis for evaluating the truth of the [testimony].'" (alterations in original) (quoting California v. Green, 399 U.S. 149, 161 (1970))).

170. *See* FRANCIS L. WELLMAN, THE ART OF CROSS-EXAMINATION 23-24 (2d ed. 1904) (discussing how to use cross-examination to discredit witnesses).

171. *See, e.g.*, Trevor J. Foster & Seth A. Northrop, *A Lawyer's Guide to Source Code Discovery*, FED. LAW., Feb. 2011, at 42, 46 (discussing some of the hidden costs of source code review in the context of civil litigation).

172. Roth, *supra* note 20, at 2040-41, 2044 (discussing the goals of reliability and dignity); *see also* David Alan Sklansky, *Confrontation and Fairness*, 45 TEX. TECH. L. REV. 103, 103 (2012) (lamenting that discussion about the Confrontation Clause "has marginalized considerations of fairness").

the eye and hear what they have to say, forcing the accuser to take responsibility in the process.[173] Reliability and accuracy may matter less here, but transparency remains key: "[T]he more inscrutable a machine process, the more its accusatory conveyances threaten the dignity of the accused and the perceived legitimacy of the process."[174] While the practical prong of confrontation aims to make machine reasoning transparent to the jury, fairness is concerned with transparency to the defendant. Still, design characteristics serve a similar role here, as they influence how effectively confrontation addresses opacity and fulfills that sense of fairness.

Accordingly, this Part considers the effectiveness of proposed forms of machine confrontation at promoting reliability and transparency for particular machine characteristics. It should be viewed as a sort of introductory guidebook about confronting technology, especially for resource-constrained defense attorneys. It considers each proposed form of confrontation, what kinds of probative value they produce, and which design characteristics they will struggle to overcome. Part III.A discusses how live testimony—though over-emphasized by modern doctrine—has real value in confronting systems involving non-quantifiable human discretion. Programmer testimony may also be a surprisingly practical substitute for source code analysis of complex codebases. Part III.B argues that source code disclosure—powerful though it may be—is likely of limited use when technology involves machine training or codebase complexity. Part III.C argues that discovery rights are particularly useful for technologies involving machine training, limited human discretion, or physical decay. Finally, Part III.D argues that "tinkering" is most useful for multi-dimensional or machine trained systems.

Subsequently, Part IV will consider whether certain characteristics require certain forms of confrontation as a matter of constitutional course. But the current Part's importance does not depend on the answer to that question. Criminal defendants and defense attorneys need an understanding of how to tackle the growing set of technologies that are swirling within the court system. This Part can also serve as a blueprint for legislative intervention, should judicial reticence continue to block the expansion of machine confrontation. Of course, this is not an exhaustive examination of how these forms of confrontation interact with different types of machines. But it should begin to illuminate the contours of the current machine landscape and firmly establish that effective confrontation cannot be machine-agnostic.

173. Roth, *supra* note 20, at 2040-41.

174. *Id.* at 2042.

A. Live Testimony

The most frequently discussed form of machine confrontation is live testimony. While live testimony may be over-emphasized in current Confrontation Clause doctrine,[175] particularly given the limitations on what it can reveal about machine reasoning, there are still instances where it offers the best form of machine confrontation. Testimony from a human operator is especially useful for machine-generated evidence involving large degrees of input transience or human discretion. Calling programmers—who are rarely the human operators—to testify may also be appropriate in cases with complex codebases or where coders make discretionary choices, such as what numerical thresholds to use.

1. Machine operators

Testimony from human operators is conventionally thought to be appropriate—and sometimes required—when the operators exercise high degrees of control or interpretation over a machine's output.[176] Where there is interpretation, such as when someone classifies a numerical prediction from a facial recognition algorithm as a "match," the machine output includes a human statement.[177] The operator will not be able to answer questions about the underlying code, for example, but they can explain how the system was calibrated, which inputs were chosen, what process was followed, and how outputs were used.[178] Operators may also be able to explain their interpretive judgments, such as in *United States v. Arce*, where an analyst had classified an image in a Cellebrite report as depicting a particular subject matter and theoretically could have later justified that classification on the stand.[179]

Conversely, it is hard to imagine that calling the forensic analyst to testify in cases like *Melendez-Diaz* would have had much practical impact because the human statement was essentially just certifying the test.[180] Even in cases

175. *See supra* Part I.A.

176. *See, e.g.*, Sites, *supra* note 51, at 101 (concluding that "it is more consistent with the Confrontation Clause's goals if courts adopt a model that preserves a right to cross-examine human operators in circumstances where the operator exercises control over the machine"); *see also supra* notes 45-47 and accompanying text.

177. *See supra* notes 45-47 and accompanying text; *cf.* People v. Goldsmith, 326 P.3d 239, 249 (Cal. 2014) (noting that a red-light camera, though it "must be programmed to activate when certain criteria are met," had no one operating it in any meaningful sense when the picture was triggered, and there accordingly was no human statement).

178. Of course, the person that interpreted the output may not be the same person that calibrated the system or inputted the data. *See* Melendez-Diaz v. Massachusetts, 557 U.S. 305, 332-34 (2009) (Kennedy, J., dissenting).

179. *See* 49 F.4th 382, 393 (4th Cir. 2022).

180. *Melendez-Diaz*, 557 U.S. at 308.

involving calibration and input selection, calling the witness to the stand may not be the most effective way to merely show that the operator followed the process; heightened reporting requirements or discovery about documentation and process is probably more useful here.[181] Discovery access to a system like COBRA, which records specific information about a breathalyzer, including individual test results, calibration information, error rates, and maintenance records,[182] is surely more useful than calling the test administrator to the stand. After all, the administrator could only testify about their imperfect memory of how the breathalyzer was used.

Human Discretion. Operator testimony is typically most helpful in cases of machines involving non-quantifiable discretion. The labeler in *Arce* made an accusation that required discretionary human judgment about the contents of the picture.[183] The law enforcement officer who pre-treats images before feeding them through facial recognition software makes similar discretionary choices. These actions are difficult to quantify and require choice from a set of legitimate options. These actions are also different in kind from the binary discretion to calibrate a breathalyzer or not. Non-quantifiable discretion is telling of the value of human testimony because it obscures understanding of the production of machine-generated evidence and because the operator is in the best position to resolve that issue.[184]

Live testimony is also helpful to persuasively question discretionary decisions before a jury, even where that discretion is quantifiable. If an officer set a grossly inappropriate threshold at which a facial recognition algorithm would find a confident match, the effect of questioning that officer in front of a jury is probably going to leave more of a lasting impression than merely introducing evidence of that threshold.

Transient Inputs. Transience may also warrant calling a human operator to testify, even where process-based discovery would otherwise be appropriate and where discretion is minimal. The person who administers a lie detector test, for example, should probably be called to the witness stand even though they are testifying about either their conclusion of the results—which does not implicate machine-generated evidence unless they explicitly cite the raw data—or about the process they employed. Because the inputs are so variable and the system so sensitive, the operator may have the best information about

181. *See generally* Cheng & Nunn, *supra* note 64 (addressing the benefits of broadened discovery rights for process-based evidence).

182. *See* Kathleen E. Watson, Note, *COBRA Data and the Right to Confront Technology Against You*, 42 N. KY. L. REV. 375, 380-81 (2015).

183. *See Arce*, 49 F.4th at 393.

184. *See* Roth, *supra* note 20, at 1979 (arguing that live testimony "should be justified based on the inability of jurors, without such testimony, to assess the black box dangers").

the specific test that took place. Discovery about the machine may also be useful. But because live testimony is familiar to attorneys and juries and fairly low cost, it may be justified here, even if its role in doctrine is overemphasized.

2. Programmers

Calling programmers to the witness stand is often rejected[185] or considered a method of last resort, largely because of the practical and epistemic challenges of deciding which programmer to call.[186] Admittedly, it is difficult to determine as a philosophical or constitutional matter who "authored" a software feature.[187] But courts are too quick to dismiss this option as impractical. It is much easier than courts think to find the author of a line of code or the expert on a particular software component. Code "repositories"—a ubiquitous tool used to maintain source code for most or all modern applications—store every file of a software application, track authorship of every line of code, and include information about how each line compares to every previous version.[188] This does not completely resolve the practical issues. In some cases, the most recent "author" may have merely added a space to a line of code that already existed. It can also be challenging to know which lines of code are relevant to the machine function being confronted. But answering this question is the bread and butter of software engineers, who do this every day. If anything, the greatest practical challenge may be figuring out the procedural mechanism for selecting and calling the appropriate

185. *See* Nguyen v. State, No. 05-20-00241-CR, 2022 WL 3714494, at *8 (Tex. Ct. App. Aug. 29, 2022) (approvingly citing previous cases that have argued that programmers have *de minimis* intervention in the final statement). *But see* People v. Wakefield, 107 N.Y.S.3d 487, 497 (N.Y. App. Div. 2019) (noting that the creator of TrueAllele "was the declarant in the epistemological, existential and legal sense"), *aff'd*, 195 N.E.3d 19 (N.Y. 2022).

186. *See* Roth, *supra* note 20, at 1986-88 (addressing the problems of calling a programmer to the stand and noting that the fact that a programmer has designed a machine does not mean "the programmer herself has borne witness to those events"); *see also* Ivan Krsul & Eugene H. Spafford, *Authorship Analysis: Identifying the Author of a Program*, 16 COMPUTS. & SEC. 233, 234 (1997) (discussing problems in determining authorship of a program like code reuse, computer formatting, and multi-person development). Machine learning, too, complicates this epistemic question. *See* Roth, *supra* note 20, at 1987.

187. This epistemic, principled problem also exists in the context of source code disclosure. For example, does the MacOS operating system need to be disclosed as part of a computer application's code? *See* Karnow, *supra* note 120, at 160-61.

188. *See* JON LOELIGER & MATTHEW MCCULLOUGH, VERSION CONTROL WITH GIT: POWERFUL TOOLS AND TECHNIQUES FOR COLLABORATIVE SOFTWARE DEVELOPMENT 34 (2d ed. 2012) (discussing one major version of this kind of source control and the fact that it "stores every version of every file").

programmer; however, evidence law already tackles this problem in other contexts, and similar methods could be employed here.[189]

Codebase Complexity. Even if identifying the right programmers is feasible, their attenuation from the final machine output often makes them poor candidates for confrontation. However, codebase complexity is the strongest indicator that live testimony from a programmer can be valuable. When code is exceedingly complex, hiring defense experts to review information will often be cost-prohibitive.[190] In expansive codebases, experts lack the synoptic view of long-enmeshed programmers, and they may not find the same unusual bugs or functional eccentricities.[191] If the goal of questioning is to understand the reasoning of the underlying algorithm, programmers may be the most efficient, centralized source of that knowledge. Programmer testimony may also be a less invasive, more focused alternative to source code disclosure which would address courts' and companies' trade secrecy concerns.[192]

Human Discretion. Like human operators, programmers often wield discretion in selecting thresholds in code,[193] so calling them to testify would help juries either understand or discredit their reasoning. In most cases, however, source code disclosure must have occurred before these thresholds are discovered, so both forms of confrontation would be needed for these machines.

Inadequate Discovery. Finally, programmer testimony may be appropriate in cases where discovery is inadequate.[194] When error rates require explanation or where simple error rates cannot readily be calculated, programmers may provide important context for a jury.

B. Source Code Disclosure

Source code disclosure is a frequently cited alternative to live testimony.[195] Disclosure is already common within civil litigation, so expanded rights in

189. Under the business records exception to the hearsay rule, for example, organizations must identify and produce a "qualified witness" with sufficient knowledge to verify the elements of the record. *See* FED. R. EVID. 803(6)(D).

190. *See* Jouvenal, *supra* note 109 (noting that defendant needed to pay $15,000 to review TrueAllele's source code).

191. *But see* Sergey Bratus, Ashlyn Lembree & Anna Shubina, *Software on the Witness Stand: What Should It Take for Us to Trust It?*, *in* TRUST AND TRUSTWORTHY COMPUTING 396, 405 (Alessandro Acquisti, Sean W. Smith & Ahmad-Reza Sadeghi eds., 2010) (arguing that programmer competency can only be judged by thorough source code review).

192. Of course, companies might argue that the discussion in front of the jury is still the main concern. If courts agree, out-of-court depositions of programmers are an alternative for complex code.

193. *See* Roth, *supra* note 20, at 1995.

194. *See infra* Part III.C.

195. *See supra* notes 88-89 and accompanying text.

criminal trials would not be wholly uncharted territory.[196] Courts also have tools like protective orders to ensure this process does not damage legitimate trade secret interests.[197] Nevertheless, judges have mostly rejected requests for source code disclosure as irrelevant or in violation of trade secret privilege.[198]

There's no doubt that source code disclosure is a crucial modern confrontation right. Disclosure, after all, provides defendants access to the actual instructions that the machine followed—a look "under the hood."[199] Barring rare transient errors where an electrical component malfunctions, algorithms merely execute the commands they are given.[200] Source code disclosure confirms the credibility of machine outputs,[201] since there are numerous places where inaccuracy and bias can enter software.[202] Indeed, source code analysis has already revealed underlying errors in cases involving the technology described in Part II.A.[203] However, it is tempting to view source code disclosure as a panacea for complex algorithmic evidence, despite the fact that it may be impractical in many individual cases.

Codebase Complexity. It is not immediately obvious whether codebase complexity weighs for or against source code disclosure. On one hand, complex codebases are more likely to have defects, whether from malice, negligence, or logical error.[204] On the other hand, there are prohibitive costs to hiring an expert to review such a complex codebase.[205] The DNA analysis tool

196. *See* Imwinkelried, *supra* note 89, at 99.

197. *See generally* Wexler, *supra* note 10 (arguing that protective orders, not a special trade secret privilege, are the appropriate tool to protect the trade secrets contained in a program's code); Chessman, *supra* note 13, at 221-22 (discussing protective orders, special independent evaluators, and in-camera review).

198. Wexler, *supra* note 10, at 1360, 1394.

199. *See* Chessman, *supra* note 13, at 183.

200. *See* Imwinkelried, *supra* note 89, at 98 ("The source code contains all the instructions that the program needs to execute its tasks."); Goutam Kumar Saha, *Transient Software Fault Tolerance Using Single-Version Algorithm*, UBIQUITY (Aug. 2005), https://perma.cc/9AAN-JPY7 (discussing transient errors and their impacts).

201. *See* Chessman, *supra* note 13, at 228 ("The only way to test the accuracy, precision, and reliability of a computer program is to see its marching orders: the source code.").

202. *See generally id.* at 186-96 (considering structural sources of error in computer code in detail).

203. *See id.* at 196-97 (discussing source code errors detected in the Alcotest 7110 and the Intoxilyzer 5000EN, two breath alcohol content testing devices employed by law enforcement in two different states).

204. *See* Hongyu Zhang, *An Investigation of the Relationships Between Lines of Code and Defects*, 2009 IEEE INT'L CONF. ON SOFTWARE MAINT. (ICSM): CONF. PROC. 274 (2009) (discussing some evidence that lines of code roughly correlate with number of defects).

205. *Cf.* Jenna Burrell, *How the Machine "Thinks": Understanding Opacity in Machine Learning Algorithms*, 3 BIG DATA & SOC'Y, Jan.-June 2016, at 4-5, https://perma.cc/Q363-N3UK ("A call for code 'audits' . . . may underestimate what this would entail as far as the

footnote continued on next page

TrueAllele, for example, uses approximately 170,000 lines of code,[206] and many other applications use millions or billions.[207] More challenging still are the applications involving labyrinthine file structure and dense, difficult-to-understand code design.[208]

It is easy to overstate these problems. Attorneys do not need to find every latent problem in a piece of software to effectively impeach the machine. In some cases, limited source code review may be cheap enough to justify the expense, even if it does not reveal anything useful. The fact that source code disclosure has been deployed in cases involving semi-complex algorithms and found bugs further validates its usefulness.[209] And lines of code can be an overstated proxy for time and cost to review, as many lines are empty space, comments from engineers, auto-generated metadata, repeated code, or otherwise irrelevant. Indeed, some public defenders recently made the choice to request—and receive—source code disclosure for TrueAllele, suggesting that they decided the cost was worth the potential benefit.[210] Still, it is hard to imagine this is an effective strategy for the majority of cases, and it seems ill-advised to suggest a public defender pursue source code disclosure as a default. Before electing to pursue source code disclosure, attorneys should consider if the codebase will likely be difficult to parse, if there is reason to think the code is unreliable, and what alternatives are available.

Machine Training. Machine learning models and other forms of artificial intelligence are poor candidates for source code disclosure. A facial recognition system is taught to detect patterns, but that decisional process is partially uncoded by design.[211] It is true that the training process (and part of the model) is represented in code, but this process is less likely to yield the decision-making reasoning or the bias, which primarily comes from the inputs

number of hours required to untangle the logic of the code within a complicated software system.").

206. State v. Pickett, 246 A.3d 279, 289 (N.J. Super. Ct. App. Div. 2021).

207. *See* David McCandless with Pearl Doughty-White & Miriam Quick, *Codebases: Millions of Lines of Code*, INFO. IS BEAUTIFUL (Sept. 24, 2015), https://perma.cc/73CM-PTJ6.

208. *See* Kroll et al., *supra* note 167, at 647 (discussing a software error that went unnoticed for years as a result of the code's nuanced structure despite public access and regular testing).

209. *See* Kirchner, *supra* note 158 (discussing errors found by a defense expert in 2016 while reviewing disclosed source code for a New York DNA testing tool); Roth, *supra* note 20, at 1994, 2024-25 (discussing errors found in STRmix, a DNA testing tool, and in Apple's "Find My iPhone" location data).

210. *Pickett*, 246 A.3d at 283.

211. *See* Kroll et al., *supra* note 167, at 638 ("[S]ource code alone teaches a reviewer very little, since the code only exposes the machine learning method used and not the data-driven decision rule.").

or user interaction.[212] For example, facial recognition is notoriously inaccurate at identifying Black or Asian faces because White subjects are overrepresented in training datasets, not because the coded model incorporates racial assumptions.[213]

Of course, source code disclosure may sometimes be the best available option for machine learning models and other forms of artificial intelligence. Testing machine correctness through other means sometimes cannot be easily employed, such as when there is no "ground truth" data for comparison[214] or when the existing validation studies do not encompass the factual background before the court.[215] When there are no easy ways to test a system externally, confrontation might require the cost of cracking it open and looking "under the hood." Even so, given the discussion above, discovery rights, programmer testimony, and tinkering are usually better alternatives for confronting machine learning models unless those options are unavailable or defense counsel can spare the extra expense.[216]

C. Broadened Discovery

Pretrial discovery is another important confrontation tool because it can reveal machine-specific information like error rates, prior statements, or internal processes. These types of information all have analogues in the human context which would be elicited through impeachment material or direct questioning.[217] Because machines cannot take the stand, these questions are replicated through discovery.

This information is not always available at the time that the machine-generated evidence is created. For example, a machine learning model may have been trained on a public dataset that is no longer available. This gap

212. *See* Ninareh Mehrabi, Fred Morstatter, Nripsuta Saxena, Kristina Lerman & Aram Galstyan, *A Survey on Bias and Fairness in Machine Learning*, 54 ACM COMPUTING SURVS. art. 115, at 4-9 (2021) (cataloguing types of bias found in machine learning models).

213. *See* Beth Findley, *Why Racial Bias Is Prevalent in Facial Recognition Technology*, JOLT DIG. (Nov. 3, 2020), https://perma.cc/5MCV-TEND. Of course, that does not mean models never incorporate racial assumptions, and Findley notes that one cause of bias is human selection of features on which to train models, which could be incorporated in code. *See id.*

214. Roth, *supra* note 23, at 222. Ground truth refers to data that are known to be correct and accurate. ENGSTROM ET AL., *supra* note 152, at 96 n.28. A human-labeled dataset of faces could be ground truth for a facial recognition algorithm.

215. *See* Imwinkelried, *supra* note 89, at 124.

216. Nevertheless, source code disclosure is still frequently discussed in the context of machine learning models, even where its limitations are recognized. *See, e.g.*, Nutter, *supra* note 12, at 939-41, 950.

217. Reputation testimony about a witness's dishonesty, for instance, is one rough analog for error rates.

cannot be reverse engineered. There are policy questions about how to handle this problem (should we incentivize good practice by declining to admit evidence absent certain discoverability?) as well as constitutional questions (is there a constitutional right to this evidence that mandates exclusion?). If a court is nevertheless willing to admit the machine-generated evidence absent this discoverable information, a different confrontation right might be appropriate. However, this Subpart assumes that this information is available.

1. Statistical Properties

One important set of information that can be disclosed through discovery is a machine's statistical properties: its error rates, training dataset, and ranges of outputs. These properties are especially useful for machine learning models. For example, they can be used to cast doubt on the reliability of a prediction or to suggest the machine's predisposition to make certain classifications, all without requiring that attorneys understand the underlying decision-making process. Still, attorneys should pay close attention to the type of model to understand the limitations of this kind of discovery.

Machine Training. Statistical properties are particularly useful for machine-generated evidence because they reveal system health without requiring a complete understanding of the system's decisionmaking rationale. First, many models already track the error rates used to optimize and train the system.[218] These values can be good proxies for accuracy. Indeed, this data probably would have been an easy way to challenge the legitimacy of the ShotSpotter report in Weeden's case.[219] Second, training data and testing data exist independently of a machine's model and are likely retained for future use and comparison. These datasets are already the source of intense scrutiny because of their capacity for bias.[220] Questioning dataset size, collection, labeling, and pre-processing can reveal a machine's lack of statistical rigor or bias.[221] Third, some models keep track of the weight they assign to certain input variables in creating predictions.[222] Because source code disclosure cannot reveal the pattern-matching that a machine learning model performs,

218. *See* GOODFELLOW ET AL., *supra* note 143, at 101-02; *see also* Nutter, *supra* note 12, at 933 ("Machine learning algorithms usually have two important error rates.").

219. *See supra* notes 1-6 and accompanying text.

220. *See, e.g.*, Lehr & Ohm, *supra* note 138, at 665 (discussing specific ways bias can enter training data, including the garbage-in-garbage-out problem); Rich, *supra* note 154, at 885 (noting that decisions about training data "allow human assumptions about what correlations *should* exist in the data to color the outcome").

221. *See* Nutter, *supra* note 12, at 935-39.

222. *See* Lehr & Ohm, *supra* note 138, at 708-10.

disclosure of this "variable importance plot" can partially substitute as a way to interrogate machine reasoning.[223]

There are, of course, limitations to each of these properties. For example, if the concern about a facial recognition system is that it disproportionately misidentifies Black faces, an attorney needs the error rate for certain racial demographics.[224] But error rates cannot always be calculated accurately.[225] To ascertain an error rate, there usually must be some "ground truth" to compare to.[226] Someone can pre-label a dataset of "drowsy" faces for an automobile drowsiness detection system, but this involves a large amount of human discretion, bias, and guesswork. In this way, a low error rate might reflect that the system is very good at identifying obviously drowsy individuals, even if both machine and labeler missed many true cases of drowsiness. Likelihood ratios are especially difficult to validate, such as with DNA testing.[227] Even where error rates are possible, they can be complicated for a jury to understand absent live testimony, particularly for machines that produce non-binary outputs.[228]

Training data has similar limitations. At a practical level, a defense attorney's presentation of bias in a dataset would not be particularly persuasive absent evidence that this bias, in fact, influenced the outcome. Moreover, examining these datasets can be challenging and costly. Analyzing demographic information in a training set is straightforward, but fruitful analysis is much harder for a human reviewing millions of unlabeled minutes of surveillance feeds for a lip-reading algorithm or raw ShotSpotter audio files. Still, finding patterns in raw data should be in the purview of defense experts, and training data is rife with potential avenues for exploration.[229]

---

223. *See id.* at 708 (naming variable importance plots as a happy medium to peek into an algorithm's reasoning).

224. *See, e.g.*, Nutter, *supra* note 12, at 934 (discussing differences in error rates across racial categories).

225. *See, e.g.*, Murphy & Rissman, *supra* note 131, at 33-34 (discussing how brain detection scans cannot easily yield error rates because it is difficult to know the base rates of "how often false or inaccurate memories happen in day-to-day life").

226. There are ways to estimate prediction quality without ground truth, but this is much less common and is the source of extensive scientific research. *See, e.g.*, Dheeraj Bhaskaruni, Fiona Patricia Moss & Chao Lan, *Estimating Prediction Qualities Without Ground Truth: A Revisit of the Reverse Testing Framework*, 2018 24TH INT'L CONF. ON PATTERN RECOGNITION (ICPR): CONF. PROC. 49.

227. *See* Roth, *supra* note 20, at 2034 (noting that "validation is a potentially incomplete method of ensuring the accuracy of machine reports in the form of statistical estimates and predictive scores").

228. For an example of the complexity of error rates in the DNA context, see generally James M. Curran, *An Introduction to Bayesian Credible Intervals for Sampling Error in DNA Profiles*, 4 LAW PROBABILITY & RISK 115 (2005).

229. *See* Nutter, *supra* note 12, at 935-39; Karnow, *supra* note 120, at 179-81.

Finally, the value of input variable maps depends on the specific type of machine learning algorithm being employed. Certain algorithms may not be able to meaningfully produce this kind of transparent mapping.[230] Or algorithms may produce misleading or uninterpretable maps because of the complex, multi-dimensional reasoning of these models.[231] To understand which statistical properties are available and helpful, attorneys—and scholars—should consider the particular machine learning models being introduced in court.

2. Process-Based Evidence

As some scholars have noted, machine-generated evidence is increasingly process-based.[232] Indeed, *Melendez-Diaz*, *Bullcoming*, and *Williams* are all examples in which most of the interpretation happened within the machine, and the analysts were largely testifying about process.[233] For these technologies, the most appropriate form of confrontation involves disclosure of "calibration results, performance reviews, standard operating procedures, company policies, design documents, and the like."[234]

These techniques are clearly amenable to at least three types of technology. First, systems with *little human discretion* depend on the process rather than the operator.[235] This kind of process is codified in operating procedures and—to an extent—design documents that specify use, inputs, and outputs. Second, systems that are at a high risk for *physical decay* are well-suited for the use of calibration results and performance reviews. Process-oriented evidence helps reveal where technology has not been adequately maintained and where decay can occur. Finally, systems with *transient inputs* like a lie detector probably warrant additional attention to process precisely because the exact input cannot be replicated.[236]

---

230. Lehr & Ohm, *supra* note 138, at 709; Will Knight, *The Dark Secret at the Heart of AI*, MIT TECH. REV. (Apr. 11, 2017), https://perma.cc/6HP4-5WLH; ENGSTROM ET AL., *supra* note 152, at 75 (discussing deep learning and other techniques that are less interpretable by design).

231. *See* Burrell, *supra* note 205, at 5, 9 (discussing the complicated interpretability problems of neural networks, which are used in handwriting analysis—one of the technologies already being used in courts).

232. *See generally* Cheng & Nunn, *supra* note 64 (addressing this trend towards process-based evidence in detail).

233. *See supra* Part I.A.

234. Cheng & Nunn, *supra* note 64, at 1106.

235. *Id.* at 1078, 1091.

236. This differs from Cheng and Nunn's rationale for process-based evidence, which points to the increased standardization of machine evidence and notes that process-based evidence is *not* merely second-best. *See id.* at 1091.

3. Prior statements, metadata, and more

The above is not comprehensive, and there may be other discovery rights that should be afforded to criminal defendants. Disclosure of prior machine statements, for example, might be required in at least two circumstances. The first is where a machine runs multiple times with a different result, suggesting the device is particularly sensitive. This discrepancy is harder to trace to a single machine design. Output variation can be the product of *transient inputs*, *machine training* that is unduly sensitive to particular features, *physical decay*, or *multi-dimensionality*. Output variations after repeated runs also occur when a human operator manipulates the input until arriving at a desired output. A straightforward conclusion here is that *highly discretionary systems* should always require disclosure of prior statements. The second circumstance in which prior statements might be required is if the machine produces multiple outputs in a single run, only some of which are inculpatory.[237] This would often be a feature of *machine-trained* algorithms that attempt to predict or classify something with some level of uncertainty and generate a list of possible matches as a result.

It may also be necessary to disclose metadata associated with machine-generated statements. For example, a red-light traffic camera would have to output the time it took the picture, the version of the onboard software, and the angle at which the camera was mounted.[238] Deepfakes—modified image, audio, or video files—are likely of minimal concern to courts for the time being, but they preview a situation where metadata could be important.[239] A recent case in the United Kingdom is the first known instance of a deepfake being introduced in court, a fact that was detected because of the audio files' metadata.[240] It is difficult to anticipate what technological characteristics warrant this kind of disclosure, though, because any output from any system could be doctored.

237. *See* People v. Knight, 130 N.Y.S.3d 919, 922 (N.Y. Sup. Ct. 2020) (discussing prosecutors' decision to disclose the top 13 possible image matches generated by a facial recognition algorithm, but to conceal the remaining 230 possible matches).

238. *Cf.* Bratus et al., *supra* note 191, at 404 (describing a case in which traffic lights were set to turn from yellow to red more quickly, manufacturing more traffic violations and allowing law enforcement to issue more tickets).

239. Pfefferkorn, *supra* note 121, at 263.

240. *Id.*

D. Tinkering

Finally, courts may permit defendants to interact with the machine accuser.[241] This could involve anything from retraining ShotSpotter's machine learning algorithm and running tests on the resulting model to getting access to the breathalyzer to be able to demonstrate how it works before the jury. As with the other confrontation rights discussed in this Part, courts would be free to impose reasonable constraints to ensure that trade secrets are protected and that the machine is not damaged.[242] This right to "tinker" with the machine could be fulfilled during discovery, but it also should grant defense counsel an opportunity to demonstrate the results before the jury.

1. Comparing against ground truth

One justification for tinkering is to verify that the machine works properly against some known samples.[243] A variation of this approach would specifically test the inputs that generated the prosecution's evidence to confirm the reported output's accuracy.[244]

Transient Inputs. If replicating the output is the goal, transient inputs limit the value of tinkering. It is possible, though, that showing that the government cannot perfectly replicate the original test from a lie detector or brain data scanner casts some doubt for the jury.[245]

Machine Training. Machine learning models might be particularly good candidates for tinkering, primarily because their opacity makes source code disclosure ineffective. This is true even if tinkering is being deployed for a different end—a lack of ground truth to compare to does not change the fact that model-based technologies need alternative confrontation avenues.

Physical Decay. Tinkering might also be one of the few ways to effectively detect a decaying system. If a defendant can retest the input and show a different output, they can demonstrate physical decay that might not even be visible through error rates.

241. *See, e.g.*, Cheng & Nunn, *supra* note 64, at 1107 (advocating for this right in the context of process-based evidence).

242. *Id.*

243. *Id.* ("If a mass spectrometer provides critical evidence in a case, the opponent may wish to test that machine using known samples.").

244. *See* Roth, *supra* note 20, at 2013-14, 2028-29.

245. Scholars have suggested that, if brain scanners are ever used in criminal trials, defendants should have access to their data. *See* Murphy & Rissman, *supra* note 131, at 47.

2. Examining trends

A second justification for tinkering is to test the reasoning of the machine by comparing trends across multiple datapoints. Properly construed, this test is about posing hypotheticals to the "witness," like attorneys do in cross-examination, rather than rigorous validation and correctness.[246]

Multi-Dimensionality. It is difficult to demonstrate faulty reasoning or find extremes when there are only one or two dimensions to vary. Given their multiple dimensions, simulations may be a good case for tinkering in this regard—particularly given their outsized impact on jurors' perceptions.[247] With so many positions and angles to change in bullet trajectory projections or airplane crash simulations, for example, defense attorneys can either highlight that the prosecution is making certain assumptions or reveal how the defense's theory of the case could be legitimate on the same set of base inputs. Drowsiness detection is another great example of the value of tinkering. It is easy to imagine facial recognition systems ascribing "drowsiness" to certain races because of a biased training set.[248] Or the system might be miscalibrated towards sudden steering movements—something that, if too sensitive, could be triggered by going over a pothole. Whatever the theory, there are multiple dimensions to consider, and access to the machine has substantial value to a defendant seeking to illuminate flaws in the evidence against them.[249]

Tinkering with a machine directly is probably less costly and time-intensive than source code disclosure in most cases (barring high training requirements to use the tool), so the absence of a particular design characteristic does not militate against its use. If protective orders and machine protection are small additional costs for the court, perhaps this right should be granted in every case. There are procedural questions about who gets to run the modified inputs and at what time, but these uncertainties can be resolved and should not prevent this beneficial confrontation right.[250]

The harder question is whether technical constraints counsel against two stronger forms of this right. First, it might be the case that certain technologies

246. *See* Roth, *supra* note 20, at 2028; Mnookin, *supra* note 95, at 578.

247. *See* Mnookin, *supra* note 95, at 576, 578; *see also* Karnow, *supra* note 120, at 157-59, 165 (discussing simulations and the numerous "unarticulated assumptions" that power them).

248. *Cf.* Adam Rose, *Are Face-Detection Cameras Racist?*, TIME (Jan. 22, 2010), https://perma.cc/QMK8-LY37 (discussing Nikon cameras that were more likely to predict that some Asian users were blinking).

249. *But see* Kroll et al., *supra* note 167, at 650-52 (discussing the process of "dynamic testing" in software engineering and cautioning that such testing cannot consider all possible inputs and does not explain *why* there is differential behavior).

250. *See* Mnookin, *supra* note 95, at 588-89.

must allow a threshold level of tinkering before they are permitted in court.[251] Courts would likely balk at requiring this threshold and excluding otherwise admissible evidence that does not allow enough input variation to test. Forcing companies to implement these changes is expensive, and it is unclear how existing programs would be treated if they fail to meet some minimum level of input manipulation.[252] There is also an administrability problem in deciding what inputs defendants can and should access, both as a constitutional matter and a practical matter.

Second, this right to confrontation by tinkering might be expanded to allow defendants to retrain machine learning models. An example would be to allow defense counsel to retrain a facial recognition algorithm with a racially diverse set of training data to show that (1) the original algorithm was biased, and (2) this bias impacted the predictive score. It is possible to show the original bias merely through the training data,[253] and this may be enough to establish doubt in the jury's mind. However, the score change would show a material impact of that bias and could strengthen reasonable doubt about the accuracy of the facial recognition system's identification. In cases where that additional layer of evidence seems likely to be determinative of the jury's ultimate verdict, attorneys may deem worthwhile the sometimes-burdensome cost of hiring an expert to retrain the model.

IV. Responding to Judicial Critiques of Practicality and Administrability

When the *Moon* court asked, "how could one cross-examine a gas chromatograph," this framework, then, is the answer.[254] How to confront a machine is primarily a practical question, one that considers the purposes that confrontation aims to achieve and the technical barriers that stand in its way. It is true—a machine cannot walk to the witness stand and verbally answer a defense attorney's questions. But those questions can be posed in the form of written interrogatories that a human handler can answer. Or those questions can be "answered" by the machine itself through error rates and output ranges. Source code disclosure can approximate answers about what factors the machine considered in its decisionmaking. Whatever the line of questioning, there are good analogues in the machine context that help achieve the core goal of the Confrontation Clause: avoiding trial by "unconfrontable but impressive-

251. Indeed, this is what Mnookin argues for in the context of introducing simulations at trial. *Id.* at 588.

252. *See id.* at 589.

253. *See supra* note 221 and accompanying text.

254. *See supra* note 66 and accompanying text.

looking" ex parte affidavit.[255] Different forms of machine confrontation are more or less appropriate for each particular machine. This discussion does not resolve all arguments against applying the Confrontation Clause to machine-generated evidence. One thing is clear, though: Machine confrontation is possible.

Of course, judges are right that there are some ways that machines are harder to confront than humans. Many of the subtle cues that attorneys watch for in cross-examination do not exist in the machine context, and some questions cannot be answered under certain machine designs. Technology evolves over time, and today's method of confrontation may not work for tomorrow's machine. The cost of machine confrontation also varies widely. Given these obstacles, it is important to depict realistically how confrontation should work, so as not to give judges additional ammunition with which to challenge doctrinal evolution.

A technology-specific approach like the one proposed here is not a legal aberration. Criminal law already makes machine-specific determinations in the context of the Fourth Amendment, for example.[256] The particularized nature of due process analysis may require similar machine-specific inquiries, especially in the face of the government's use of increasingly opaque artificial intelligence.[257] And even if machine confrontation is truly distinct from the way we treat human witnesses, the important rights at issue here and the unique nature of machine witnesses should justify the departure from precedent.[258] If anything, perhaps the law should treat human witnesses in the same way.[259]

A harder question is whether the practical considerations of Part III suggest that certain forms of confrontation are constitutionally mandated or, alternatively, constitutionally inadequate. Source code disclosure, for example, does no more than live testimony to protect against the "unconfrontable but impressive-looking" facial recognition machine learning model, which forms decisionmaking patterns not visible from code.[260] In this sense, source code disclosure alone may be inadequate as a constitutional protection for machine

255. Roth, *supra* note 20, at 2041; *see supra* notes 69-70 and accompanying text.

256. *See* Sites, *supra* note 15, at 572 (considering the Fourth Amendment's evolution to include modern technology).

257. *See* ENGSTROM ET AL., *supra* note 152, at 82-85.

258. *Cf.* Mnookin & Kaye, *supra* note 94, at 103 (advocating for a similar form of Confrontation Clause exceptionalism in the context of scientific experts). *But see* Sites, *supra* note 51, at 91 (noting that "the Supreme Court has cautioned against Confrontation Clause tests that are too malleable").

259. *See generally* Roth, *supra* note 23, at 223 (arguing that a broader view of confrontation would and should extend beyond machine conveyances, including to human witnesses).

260. Roth, *supra* note 20, at 2041; *see supra* notes 211-12 and accompanying text.

learning models. This is not an easy issue, and it requires more consideration and a deeper parsing of constitutional text than this Note can provide. But the source code example suggests that some forms of machine confrontation may fail to clear a constitutional floor, and others may be required under the Confrontation Clause.

As a matter of constitutional line drawing, a technology-specific approach is admittedly a difficult standard to administer. The list of technologies is long and dynamic, meaning that this inquiry must largely proceed on a time-consuming and fuzzy case-by-case basis. The *Melendez-Diaz* dissent concerns come roaring back as well.[261] If calling a human witness is a matter of constitutional concern, how do we decide between the operators, programmers, data scientists, and contractors? What about the labelers that interact with machine learning datasets? Stepping away from human testimony, how do we deal with cases where no single form of confrontation is adequate? Does the Constitution demand a combination of approaches?

These questions add complexity, but they are probably overstated. As mentioned above, criminal law already makes these kinds of case-by-case determinations, even in the context of trial proceedings.[262] Similarly, judges generally have discretion to *limit* certain lines of cross-examination, even though these ad hoc determinations can rise to the level of Sixth Amendment violations.[263] Why can courts not *require* lines of confrontation for a particular machine? These in-court determinations might be difficult for unfamiliar algorithms, but the median case can probably be resolved quickly.

Ultimately, the technology-specific approach proposed here still matters, regardless of how the Confrontation Clause grows to accommodate machine-generated evidence. First and foremost, defendants and defense attorneys can use the terminology and technical understanding in Parts II and III to "confront" machines, even if they only do so during authentication or other evidentiary proceedings. As these machines become increasingly common, this understanding is crucial.

Second, even if courts reject the notion that machines are constitutionally confrontable, legislatures can carry the torch towards better evidentiary guarantees. Many policy levers are available, some of which have already been discussed in scholarship and some of which warrant much deeper discussion.[264] The Federal Rules of Evidence could be amended to place stronger authentication or expert testimony requirements on certain types of

261. *See supra* notes 38-39 and accompanying text.

262. *See supra* note 256 and accompanying text.

263. Roth, *supra* note 20, at 2051.

264. For some further discussion, see, for example, Murphy, note 99 above, at 777-97.

algorithms.[265] Statutory protections could provide enhanced pretrial discovery or eliminate trade secret privilege, promoting two of the proposed forms of confrontation discussed in this Note.[266] Legislatures could also fund or promote open-source software to avoid the costs and litigation of source code disclosure.[267] Whatever the solution, policymakers must understand how the type of technology impacts the effectiveness of the proposed procedural right.

Conclusion

Something is broken in our evidentiary approach to machines. For decades now, judges, jurists, and scholars have spoken in platitudes, treating self-driving cars, thermometers, gunshot detectors, and breathalyzers as if they are one and the same. Much attention has been paid to the way jurisprudence carves out accusations *by machines* as a class, as if machines lack the same capacity for bias, opacity, and inaccuracy as humans. That attention is warranted. But the devil is in the details. Judges have dismissed machine confrontation in part because of the lack of clarity about how it would work in practice. Machine confrontation desperately needs that clarity now, lest efforts to jumpstart Confrontation Clause doctrine fall to a meaningless, underbaked set of procedural rights. And lawyers, judges, and legislatures alike benefit from the practical framework described here, even if the Confrontation Clause does not become the vehicle to incorporate it. Hopefully something changes before the next wave of sophisticated, inculpatory technology enters the courtroom.

265. *See, e.g.*, Roth, *supra* note 23, at 222-23 (discussing modifying rules around expert testimony). For extended discussion on reforming other evidentiary rules, see Roth, note 20 above, at 2022-39 (hearsay); and Brian Sites, *Machines Ascendant: Robots and the Rules of Evidence*, 3 GEO. L. TECH. REV. 1, 15-27 (2018) (hearsay and authentication).

266. Roth, *supra* note 23, at 223. Indeed, Roth discusses some legislation that has already been pursued in California. *Id.*

267. Chessman, *supra* note 13, at 223-24.