Automating Junk Science

LISA WATERS[†]

In the late twentieth century, the boom of forensic disciplines in criminal prosecutions helped drive mass incarceration to an all-time high. Yet scientific and legal inquiry revealed a disturbing truth: Most forensic methods accepted in criminal courts are entirely lacking in empirical support or scientific foundation—in other words, "junk science." Forensic proponents have recently turned to computer algorithms, costly equipment, and proprietary trade secrets litigation to defend dubious techniques, ushering in a second wave of forensic reliance. But automated technology has masked rather than cured the foundational infirmities in these forensic fields.

This Article examines the growing trend of automated forensics using firearms examination as a case study. The Article demonstrates that firearms examination meets the criteria for junk science, then shows how its proponents are turning to virtual imaging and comparison technology in an attempt to rehabilitate the discipline. But automation fails to correct for underlying scientific shortcomings, instead obscuring them while adding further potential for mechanical and algorithmic error.

Troublingly, the turn to algorithms in defending firearms examination is far from an isolated strategy. Similar efforts are seen with policing tools such as facial recognition technology, DNA analysis software, and risk assessment algorithms. This Article illustrates how the traditional means of challenging forensic evidence with case-by-case litigation are ill-equipped to address the rise in automated forensic technology and examines the risk of automation reanimating and perpetuating the injustices borne of our troubled forensic history. The Article contends that we must eliminate junk science from criminal courtrooms, while also proposing an intermediate path forward through reforms that prioritize transparency, meaningful scientific and legal scrutiny, and the rights of the criminally accused.

[†] Associate Professor of Law, City University of New York (CUNY) School of Law. Former Felony Trial Attorney, Managing Attorney, and Forensic Science Working Group Member at the New Jersey Office of the Public Defender. Many thanks to Amanda Savage, Nina Chernoff, Daniel Loehr, Babe Howell, Nancy Leong, Christopher Lau, Maneka Sinha, John Wagner, Richard Guttierez, Sarah Gottlieb, Ngozi Ndulue, Anna Roberts, Eve Hanan, Asees Bhasin, and other participants at the AALS Clinical Conference and CrimFest who provided valuable feedback and advice. I am grateful to David De La Cruz, Lou Kreitler, and Morgan Bissett-Tessier for excellent research assistance and to the editors of the *UC Law Journal* for their improvements to this Article. Thanks also to Bob Miseo, Liz Jarit, and David Ghigliotty.

TABLE OF CONTENTS

INTRODUCTION

In June of 2004, T.P.[1] was found alongside a New Jersey highway with a gunshot wound to his head.[2] He was rushed to the hospital but died within hours. The local prosecutor's office quickly launched a homicide investigation. Detectives interviewed family members, scoured the crime scene, and collected the fatal bullets as evidence.[3] But no arrests were made, no charges filed—and for good reason: There was simply no evidence pointing to a suspect.

The next year, a man named David Ghigliotty[4] was arrested for possessing a gun without the proper permit. Because the gun was loosely connected with T.P., a friend of Mr. Ghigliotty, the prosecutor's office hired a firearm examiner to compare the gun with the homicide crime scene bullets.[5] The examiner, employing the same method of forensic firearms testing used for well over a century, compared the bullets under a microscope and concluded that Mr. Ghigliotty's gun did *not* fire the fatal bullets.[6] The case of T.P.'s homicide remained as before—unsolved. No arrests were made, no charges filed.

That changed over a decade later. In 2015, a detective at the prosecutor's office began reviewing the unsolved T.P. homicide file and asked their new in-house firearm examiner, Lieutenant Michael Sandford, to re-examine the bullets. Sandford ultimately agreed.[7] Over the course of several months, Sandford compared bullets from T.P.'s homicide with bullets fired from Mr. Ghigliotty's gun in 2005.[8] These examinations were both undocumented and fruitless; Sandford was unable to reach any conclusions.[9]

Then Sandford learned about BULLETTRAX. While at the 2016 annual conference of the Association of Firearm and Toolmark Examiners ("AFTE"),

---

1. Although the decedent's full name is publicly available in the published opinion in *State v. Ghigliotty*, 232 A.3d 468, 474 (N.J. Super. Ct. App. Div. 2020), this Article uses only initials to preserve his privacy.

2. *See id.*

3. *Id.* ("The police did not recover a murder weapon at the scene but were able to retrieve three bullets from the victim's body. . . . [However, o]nly one of the projectiles recovered at the murder scene, a bullet jacket fragment that was 'significantly damaged,' was deemed intact enough to be suitable for comparison.").

4. Unlike the decedent, this Article uses Mr. Ghigliotty's full name, as it appears in the published court opinion, which cannot be cited without using his last name. *Id.* While I would prefer to maintain some anonymity for Mr. Ghigliotty, his identity is central to the legal proceedings and already part of the public record through official court documentation. This Article contains no confidential or privileged information.

5. *See id.* (explaining that in 2005, police arrested Mr. Ghigliotty for the unlawful possession of a handgun without a permit, and records revealed that T.P.'s brother had purchased the handgun in October 2003).

6. *See id.* at 474–75 (describing how a forensic firearm examiner used standard comparison microscopy to compare the bullet jacket fragment from the homicide with test shots from this handgun and concluded that they were not a match—finding "negative results" or "an elimination" that established no connection between the fragment and the weapon).

7. *See id.* The prosecution's decision to reopen the case over a decade later was notable given that Sandford initially resisted re-examining another examiner's work, stating it violated standard practice. Only after repeated requests did Sandford agree to conduct the re-examination that had previously yielded an elimination.

8. *See id.* at 474–76. Notably, the details and dates of these examinations were not documented.

9. Despite extensive re-examination using traditional comparison microscopy—including taking approximately ten additional test shots from different ammunition manufacturers and reviewing the evidence "many" times—Sandford testified that "he still could not reach a conclusion on the question of whether the fragment could have come from defendant's handgun." *Id.* at 475–76.

he attended a workshop on emerging technology purporting to advance the capacity, efficiency, and accuracy of firearms examinations—3D scanning and virtual comparison microscopy ("VCM")—and in particular a tool called "BULLETTRAX," then a product of Ultra Electronics Forensic Technologies.[10] Sandford arranged a visit to their laboratory in Montreal. He brought evidence bullets from the T.P. file and used BULLETTRAX to create computer-generated images of these bullets using a proprietary, unregulated algorithm, and a companion software called Matchpoint to virtually modify and examine the images. Using these virtual renderings "like a GPS," Sandford found his way to a new conclusion,[11] at odds with the findings of all previous examinations, including his own. He concluded that the gun seized from David Ghigliotty in 2005 did, in fact, fire the bullets that killed T.P.[12]

The local prosecutors had never heard of BULLETTRAX, yet alone encountered it in a case. But they readily deferred to this algorithm-aided conclusion over the contrary findings of their first expert. After all, this was a computer! Surely reliable, surely legitimate. Murder charges were filed. Mr. Ghigliotty was arrested, detained, and faced a sentence of life in prison.[13]

Then a trial attorney at the New Jersey Office of the Public Defender, I joined Mr. Ghigliotty's defense team a year after his arrest. I, like my colleagues, the prosecutor, and judge alike, had never heard of BULLETTRAX. I dove into the discovery (which provided little information) and researched the technology. What I found was a dearth of available literature—no academic articles, no published validation studies, no defense attorneys that had encountered the tool in practice, no federal regulations, standards, or rules governing use of the technology, and, most frustratingly, no defense-oriented firearm experts with subject matter knowledge about this emerging technology.

The defense team filed a motion challenging admissibility of the expert conclusion relying on BULLETTRAX and demanding disclosure of the proprietary software. The trial court scheduled a *Frye* hearing[14] and granted the discovery request. Following interlocutory appeal, the latter order was reversed and remanded as premature.[15] The appellate court insisted on a more "definitive

10. BULLETTRAX was described as using "computer automation and confocal microscopy to scan the surface of a bullet, read its topography in 3D, and create 2D and 3D images," allowing examiners to "compare up to six images at one time on a single screen as opposed to a single pair" with advantages over traditional comparison microscopy. *Id.* at 476.

11. When Sandford testified that he used the screen shots from BULLETTRAX "like[] a GPS" to guide his comparison microscope analysis of the "areas of interest" and "came to an opinion of an identification or a positive identification." *Id.* at 477. Sandford acknowledged that he was "not able to make a positive identification prior to using the BULLETTRAX technology." *Id.* at 478. However, he claimed his conclusions were based on his "live comparison on a comparison microscope" rather than "photographs or a computer image." *Id.* at 477–78.

12. *Id.* at 478–79.

13. The penalty for murder in New Jersey ranges from thirty years to life in prison. N.J. STAT. ANN. § 2C:11-3 (West 2024).

14. Hearing to determine admissibility of forensic evidence under *Frye v. United States*, 293 F. 1013, 1014 (D.C. Cir. 1923); *Ghigliotty*, 232 A.3d at 479.

15. *Ghigliotty*, 232 A.3d at 485–86.

showing of . . . need"[16] before mandating disclosure of the proprietary software and specifically noted the absence of input from a defense expert.[17] In the context of obstacles encountered in advancing the litigation, the decision highlighted ways in which the criminal legal system is simply ill-equipped to address and appropriately regulate the use of proprietary technology in criminal cases.

Soon after it was remanded for further fact-finding, the case resolved, prematurely terminating litigation around both disclosure and admissibility of evidence relying on BULLETTRAX and Matchpoint methods. Years later, the legal questions raised in *Ghigliotty* remain unanswered, and the academic literature on automated firearms examination tools remains remarkably thin.

BULLETTRAX is one tool within a field of emerging forensic technology, which purports to modernize[18] forensic firearms examinations by using proprietary algorithms. *Ghigliotty* featured the first known[19] instance of the technology being challenged through direct litigation. 3D scanners and VCM have gained attention[20] and usage[21] by firearms examiners in the years since *Ghigliotty*, yet, troublingly, expansion has not been paralleled with meaningful response by the legal community. The lack of transparency, access, and meaningful opportunity for review evidenced in *Ghigliotty* will only compound as usage of the technology evolves, erecting substantial barriers to ensuring the integrity of forensic evidence and the constitutional rights of the accused.

Moreover, automated firearms examination is merely one illustration of a broader trend. By minimizing the role of the human examiner and relying

16. *Id.* at 486. The appellate court held that a "defendant is required to make a more definitive showing of his need for this material to provide the court with a rational basis to order the State to attempt to produce it," finding the trial court's discovery order was premature because "the defense did not present a certification from an expert in support of this claim for disclosure." *Id.*

17. *Id.* Notably, the defense team was unable to locate an expert with both the computer science background and subject matter knowledge to assist this particular request; however, a better record would have included input from an expert with computer science expertise and/or expertise in analogous forensic tools and/or documentation in the record of the efforts made to identify and retain an expert, without success.

18. VCM has actually been used for decades by law enforcement ballistic databases, while remaining hidden in the "shadows" of the investigative stage. *See National Integrated Ballistic Information Network (NIBIN)*, BUREAU OF ALCOHOL, TOBACCO, FIREARMS & EXPLOSIVES (Jan. 28, 2026), https://www.atf.gov/firearms/national-integrated-ballistic-information-network-nibin; Christopher Lau, *Shadow Forensics: Uncovering 911 Call Analysis*, 111 CORN. L. REV. (forthcoming 2025) (manuscript at 25) ("Shadow Forensics is the process by which the State effectively elides gatekeeping: these forensics are not proffered, are not challenged, and therefore escape any evidentiary examination by a judge before being presented as expertise to fact finders in pretrial hearings or, more rarely, at trial."). Only recently, however, has the technology taken a more central and pervasive role in direct case work, as illustrated in *Ghigliotty.*

19. *See Ghigliotty*, 232 A.3d at 477 n.7. Nationwide inquiries revealed no other instances in 2020. VCM has been challenged in other cases since *Ghigliotty*, with similar results—premature termination of litigation due to case resolutions.

20. Every year the Association of Firearm and Toolmark Examiners Annual Conference features presentations on VCM. *See AFTE 2025: Workshops*, ASSOC. FIREARM & TOOLMARK EXAM'RS, https://na.eventscloud.com/website/74655/workshops (last visited June 12, 2025); *AFTE 2025: Technical Session*, ASSOC. FIREARM & TOOLMARK EXAM'RS, https://na.eventscloud.com/website/74655/technical (last visited June 12, 2025).

21. There are known recent instances of its use in New York, North Carolina, South Carolina, Oregon, Georgia, and New Mexico. This information has been gathered through conversations and nationwide inquiries of defender and forensic communities; it is by no means exhaustive.

heavily on proprietary computer programs, VCM joins the movement of automated policing throughout the criminal legal system, sharing with that trend a number of risks. Automated technology has touched all stages of policing and prosecution, from facial recognition technology used to identify suspects,[22] to risk assessment tools relied on by judges for bail and sentencing decisions,[23] to DNA genotyping software purporting to parse and compare complex DNA mixtures accurately.[24] These tools present unique issues, risks, and barriers to meaningful and fair confrontation, yet they remain unregulated, non-transparent, and poorly addressed by legal constructs designed to govern a different era of forensic evidence.

Problems with forensic evidence, however, existed long before automation. Most forensic methods were developed in police departments to pursue investigations and convictions.[25] Despite adopting the terminology of science, "other than DNA analysis, forensic disciplines did not arise out of academia, research institutions, or scientific laboratories—they do not have their origins in the sciences at all."[26] These origins lead to forensic evidence's natural alignment with law enforcement and prosecutorial interests and adoption of a "carceral culture,"[27] prioritizing outcomes and resisting review, reform, and transparency efforts. Automated forensic tools follow the same patterns of development and culture. They are created for law enforcement uses and are implemented by private companies with commercial incentives to cater to carceral priorities.[28]

22. *See, e.g.*, ELIZABETH F. LOFTUS, JAMES M. DOYLE, JENNIFER E. DYSART & KAREN A. NEWIRTH, EYEWITNESS TESTIMONY: CIVIL AND CRIMINAL § 1-6[b] (6th ed. 2025); Jennifer Lynch, *Face Off: Law Enforcement Use of Face Recognition Technology*, ELEC. FRONTIER FOUND. (Apr. 20, 2020), https://www.eff.org/wp/law-enforcement-use-face-recognition; Kate Conger, Richard Fausset & Serge F. Kovaleski, *San Francisco Bans Facial Recognition Technology*, N.Y. TIMES (May 14, 2020), https://www.nytimes.com/2019/05/14/us/facial-recognition-ban-san-francisco.html; Joseph Cox, *ICE Spends Millions on Clearview AI Racial Recognition to Find People 'Assaulting' Officers*, 404 MEDIA (Sep. 8, 2025, at 11:48 ET), https://www.404media.co/ice-spends-millions-on-clearview-ai-face-recognition-to-find-people-assaulting-officers.

23. *See, e.g.*, Jeff Larson, Surya Mattu, Lauren Kirchner & Julia Angwin, *How We Analyzed the COMPAS Recidivism Algorithm*, PROPUBLICA (May 23, 2016), https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm (discussing their analysis of the Correctional Offender Management Profiling for Alternative Sanctions recidivism algorithm).

24. *See* Katherine L. Moss, Note, *The Admissibility of TrueAllele: A Computerized DNA Interpretation System*, 72 WASH. & LEE L. REV. 1033, 1060 (2015).

25. *See* NAT'L RSCH. COUNCIL, STRENGTHENING FORENSIC SCIENCE IN THE UNITED STATES: A PATH FORWARD 42, 187 (2009), https://www.ojp.gov/pdffiles1/nij/grants/228091.pdf [hereinafter NRC REPORT]; *see also* Maneka Sinha, *Radically Reimagining Forensic Evidence*, 73 ALA. L. REV. 879, 894 (2022) (discussing the origins of forensic methods).

26. Sinha, *supra* note 25; *see* Eric S. Lander, *Fixing Rule 702: The PCAST Report and Steps to Ensure the Reliability of Forensic Feature-Comparison Methods in the Criminal Courts*, 86 FORDHAM L. REV. 1661, 1668 (2018); *see also* Paul C. Giannelli, *Independent Crime Laboratories: The Problem of Motivational and Cognitive Bias*, 2010 UTAH L. REV. 247, 250 (2010) (discussing the potential dangers of having biases in independent laboratories); NRC REPORT, *supra* note 25, at 42 (explaining that many forensic tests have "never been exposed to stringent scientific scrutiny"); Radley Balko, Opinion, *Jeff Sessions Wants to Keep Forensics in the Dark Ages*, WASH. POST (Apr. 11, 2017), https://www.washingtonpost.com/news/the-watch/wp/2017/04/11/jeff-sessions-wants-to-keep-forensics-in-the-dark-ages (criticizing Jeff Sessions's approach to forensics).

27. Sinha, *supra* note 25, at 898.

28. *Infra* Subpart II.E.

Unsurprisingly, these tools replicate underlying scientific shortcomings, while adding novel concerns unique to the nature of automation.

Automating forensics in criminal case work raises new and important questions for a legal community that has yet to reckon with the prevalence and harm of junk science in criminal courtrooms. How do we resolve tensions between constitutional rights and trade secret rights? How do legal rules contemplating human experts apply to automated machines? What role should science play in criminal courts? These and other questions remain unanswered by the rules and procedures of criminal law.

Scientifically unsound forensic evidence—automated and otherwise—remains ubiquitous in modern prosecutions. The harm of junk science in criminal courts can only truly be solved by disclaiming these forensic disciplines entirely. Tailored legal rules and reforms will not cure the scientific deficiencies in forensic methods, but they can work to unveil the flaws that are increasingly obscured by algorithms so that they can be meaningfully scrutinized and challenged on both case-by-case and systemic levels. Recognizing the need for "non-reformist reforms"[29] to protect those most impacted while awaiting a more radical reordering of values, this Article suggests concrete reforms—novel approaches to novel issues—to more appropriately address, gatekeep, and scrutinize the reliability of automated forensic evidence and safeguard the rights of the criminally accused.

This Article proceeds as follows. Part I defines and explores "automated junk science" and its impact on criminal prosecutions by taking a closer look at the technology relied upon in *Ghigliotty*, 3D scanners and virtual comparison microscopy, and the underlying discipline of forensic firearms examination. Part II examines the impact of automation on forensic examinations, considering the ways that algorithms change the *procedure* of forensic analyses and introduce novel issues of scientific and legal import, while also failing to address the *nature* of scientifically deficient forensic methods. Automation does not cure the fatal infirmities of junk science; rather, it both adds and obscures them with a façade of tech legitimacy, threatening to dissuade scrutiny and undo progress in uncovering the flawed nature of forensic evidence. Part III turns to the criminal legal system, documenting the ways in which it both structurally and procedurally fails to account for the novel issues raised by automation. Considering historical failures of the criminal legal system to screen out junk science, contributing to countless wrongful convictions, Part IV suggests systemic and rule changes to address risks raised by the entrance of automation into the already problematic forensic toolkit of law enforcement, urging the

29. *See* Dan Berger, Mariame Kaba & David Stein, *What Abolitionists Do*, JACOBIN (Aug. 24, 2017), https://jacobin.com/2017/08/prison-abolition-reform-mass-incarceration ("Central to abolitionist work are the many fights for non-reformist reforms—those measures that reduce the power of an oppressive system while illuminating the system's inability to solve the crises it creates."); Sinha, s*upra* note 25, at 892 ("A non-reformist approach prioritizes contraction of the carceral institution and its power and avoids actions that validate or condone the current system.").

abolition of junk science, while also providing a blueprint for intermediate reforms.

I. FORENSIC FIREARMS EXAMINATION: A CASE STUDY OF AUTOMATED JUNK SCIENCE

Forensic methods have long been used to surveil, prosecute, convict, and punish—the "core inputs and outputs of the criminal legal system."[30] Pervasive in criminal case work, forensic methods appear in cases as minor as shoplifting and as serious as murder. Yet forensic evidence has proven to be as unreliable as it is ubiquitous.[31] Following a slew of exonerations in convictions relying on forensic evidence,[32] committees of scientists and experts convened to examine what was going so wrong with forensics.[33] Each inquiry revealed the same simple truth: The vast majority of forensic "sciences" are simply not science.[34] Developed in "police departments as investigative aids meant to produce evidence that would connect suspects to crimes and secure convictions,"[35] it is no surprise that forensic evidence lacks the features of real science.[36]

After a century as a fixture in criminal investigations, scientific inquiry revealed the majority of forensic disciplines to be "junk science."[37] Some

30. Sinha, *supra* note 25, at 892.

31. Forensic evidence plays a significant role in criminal proceedings, though measuring its exact prevalence presents challenges due to the vast number of cases and varying evidentiary standards across jurisdictions. NRC REPORT, *supra* note 25, at 95. Despite these measurement difficulties, the importance of forensic evidence is reflected in juror expectations, with scientific evidence of some kind expected by 46% of jurors in every case, rising to 74% in murder cases and 73% in rape cases. *See Percentage of Jurors Who Expect Scientific Evidence from Prosecution*, NAT'L INST. OF JUST. (Mar. 2008), https://nij.ojp.gov/media/image/19656.

32. *See Misapplication of Forensic Science*, THE INNOCENCE PROJECT, https://innocenceproject.org/misapplication-of-forensic-science (last visited June 23, 2025) (explaining forensic evidence played a role in convicting over half of the Innocence Project's DNA exonerated clients); *see also* Vanessa Meterko, *Strengths and Limitations of Forensic Science: What DNA Exonerations Have Taught Us and Where to Go From Here*, 119 W. VA. L. REV. 639, 640 (2017) (finding that incorrect forensic conclusions contributed to nearly half of 343 DNA exonerations cases examined); Emily West & Vanessa Materko, *Innocence Project: DNA Exonerations, 1989-2014: Review of Data and Findings From the First 25 Years*, 79 ALB. L. REV. 717, 743 (2016) (finding that the misapplication of forensic science played a role in convicting 47% of the 325 DNA exonerated cases the authors analyzed).

33. *See generally* NRC Report, *supra* note 25; PRESIDENT'S COUNCIL ADVISORS ON SCI. & TECH., EXEC. OFF. PRESIDENT, FORENSIC SCIENCE IN CRIMINAL COURTS: ENSURING SCIENTIFIC VALIDITY OF FEATURE-COMPARISON METHODS (2016), https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_forensic_science_report_final.pdf [hereinafter PCAST REPORT].

34. *See generally* PCAST REPORT, *supra* note 33 (finding pattern-matching forensic fields, including firearm and toolmark examination, lacking in foundational scientific validity).

35. Sinha, s*upra* note 25, at 894.

36. *See* PCAST REPORT, *supra* note 33, at 46 ("[T]he fundamental principles of the 'scientific method'—applicable throughout science—that valid scientific knowledge can *only* be gained through *empirical* testing of specific propositions."); *Scientific Method*, OXFORD ENG. DICTIONARY ONLINE (Mar. 2014), https://www.oed.com/dictionary/scientific-method_n?tab=meaning_and_use (Defining the scientific method as a "method or procedure . . . consisting in systematic observation, measurement, and experimentation, and the formulation, testing, and modification of hypotheses.").

37. Modern standards for the admission of scientific evidence in federal court are governed by Rule 702, which is followed by many, but not all, states. Rule 702, modified after the Supreme Court's ruling in *Daubert v. Merrell Dow Pharms., Inc.*, 509 U.S. 579 (1993), requires, among other criteria, that the expert testimony: "is

forensic practitioners responded to the scientific critiques by discounting or rejecting them.[38] Others turned to technology. This Part explores the automation of junk science and its impact on criminal cases by taking a closer look at the forensic discipline at issue in *Ghigliotty*: firearms examination.

A. FORENSIC FIREARMS EXAMINATION: A JUNK SCIENCE

Forensic firearms examination is junk science.[39] Scientists and independent experts have concluded there is inadequate empirical evidence to support even the most basic premise of the discipline: that a single firearm will consistently produce unique markings.[40] Yet firearm examiners have become a centerpiece of criminal prosecutions,[41] and when forensic expert testimony is admitted at trial, as it almost always is,[42] jurors tend to believe it.[43]

---

based on sufficient facts or data"; "is the product of reliable principles and methods"; and "reflects a reliable application of the principles and methods to the facts of the case." FED. R. EVID. 702; *Daubert*, 509 U.S. at 597. Forensic evidence that would not meet the admissibility test (if properly applied) laid out in Rule 702 is referred to throughout this article as "junk science." The term has been adopted by Supreme Court Justices, *General Electric Company v. Joiner*, 522 U.S. 136, 153 (1997) (Stevens, J., concurring in part and dissenting in part), academic discourse, Aliza B. Kaplan & Janis C. Puracal, *It's Not a Match: Why the Law Can't Let Go of Junk Science*, 81 ALB. L. REV. 895 (2018), and the media, Liliana Segura & Jordan Smith, *Bad Evidence: Ten Years After a Landmark Study Blew the Whistle on Junk Science, the Fight Over Forensics Rages On*, THE INTERCEPT (May 5, 2019), https://theintercept.com/2019/05/05/forensic-evidence-aafs-junk-science.

38. *See* Radley Balko, *Devil in the Grooves: The Case Against Forensic Firearms Analysis*, WATCH (May 25, 2023), https://radleybalko.substack.com/p/devil-in-the-grooves-the-case-against. ("[The firearm examiners community] and their defenders in law enforcement and prosecutors' offices have generally responded to such criticism not with humility or regret, but by pushing back, sometimes belligerently and dishonestly. And unfortunately, the courts have mostly followed their lead."). There are notable exceptions, particularly in the field of bite mark comparison. Doctors Iain Pretty and Adam Freeman both once championed the discipline of bite mark comparison and testified as prosecution experts. Alarmed by the results of the 2009 NRC Report, they began their own scientific testing, some including their own case files. Their results were alarming; in nearly all cases, participants (all board-certified dentists) could not even agree on whether they were looking at a bite mark at all! Doctors Pretty, Freeman, and many other forensic dentists have now rejected bite mark comparison as foundationally unsound. *See* Daniele Selby, *Why Bite Mark Evidence Should Never Be Used in Criminal Trials*, THE INNOCENCE PROJECT (Apr. 26, 2020), https://innocenceproject.org/news/why-bite-mark-evidence-should-never-be-used-in-criminal-trials. Doctor Freeman now freely offers his expertise to educate courts on the lack of foundational validity of bite mark comparison, and aid efforts to exonerate those falsely convicted based on bite mark evidence.

39. *See infra* Subpart.I.B.

40. *See* NAT'L RSCH. COUNCIL, BALLISTIC IMAGING 3 (2008) [hereinafter BALLISTIC IMAGING].

41. The expansion of firearms examination was fostered by the explosion of federal funding to law enforcement throughout the mid-twentieth century, fueling the creation, expansion, and greater use of forensic laboratories solely for law enforcement purposes *See, e.g.*, Sinha, *supra* note 25, at 894–95 (discussing how most forensic methods were developed in police departments, not scientific institutions, and were fueled by federal initiatives such as the War on Crime and the Law Enforcement Assistance Act).

42. *See* Brandon L. Garrett, Eric Tucker & Nicholas Scurich, *Judging Firearms Evidence*, 97 S. CAL. L. REV. 101, 124 (2024) (explaining that by the 1970s and 1980s, courts routinely admitted firearms expert testimony without discussion); *see, e.g.*, Hampton v. People, 465 P.2d 394, 400 (Colo. 1970) (stating there was no abuse of discretion for admitting a firearm comparison expert's testimony). For perhaps the first case referring to the discipline as a type of toolmark comparison, see United States v. Bowers, 534 F.2d 186, 193 (9th Cir. 1976).

43. Brandon L. Garrett, Nicholas Scurich & William E. Crozier, *Mock Jurors' Evaluation of Firearm Examiner Testimony*, 44 LAW & HUM. BEHAV. 412, 416–18, 422 (2020) (concluding that jurors place significant weight on firearm examiner testimony declaring a match regardless of the specific language used); *id.*

This series of events—expansion of forensics in criminal investigations, lax judicial gatekeeping, and juror deference—helped enable incalculable wrongful convictions, with the misapplication of forensic evidence playing a role in convicting over half of the Innocence Project's DNA exonerated clients.[44] Because courts are reluctant to revisit convictions absent DNA evidence, these numbers fail to grasp the breadth of wrongful convictions involving firearms analysis.[45]

There have been a number of high-profile exonerations involving firearm toolmark evidence, including Patrick Pursley in Illinois[46] and Anthony Ray Hinton in Alabama.[47] The well-publicized trials of Curtis Flowers in Mississippi[48] relied on firearm toolmark evidence in addition to flawed eyewitness identification evidence, another common feature of wrongful convictions. Despite dire warnings broadcast by these exonerations, firearm experts continue to routinely testify in criminal trials. The unfounded deference afforded to forensic firearm experts by courts and jurors may result, in part, from a lack of information about the discipline's features, methods, and deficiencies.

1. Firearms Examination 101

Firearms examination is the most common sub-category of a broader forensic discipline called toolmark examination. A "toolmark" is an impression left by the contact of a "tool" (any object) with a softer surface or "work piece."[49] In firearms examinations, the "tool" is the firearm, and the "work pieces" are discharged bullets and/or cartridge cases[50] (often described as "projectiles" or

(suggesting that jurors are more likely to convict when a "match" is declared, which indicates that they often accept such testimony at face value).

44. *See* THE INNOCENCE PROJECT, *supra* note 32.

45. *See* Balko, *supra* note 38. The reluctance of courts to revisit convictions absent DNA evidence is particularly challenging for cases involving firearm evidence, compared with those involving inherently biological forensic evidence. Other faulty forensic fields like bite mark comparison, hair analysis, or fingerprint identification are biological in nature, meaning "evidence that may have been incorrectly matched is likely to include or have been accompanied by DNA. So DNA tests can confirm or refute the analyst's conclusions." *Id.* Bullets are not biological; typically, a shooter's DNA is not left alongside ballistic evidence.

46. *See Patrick Pursley*, BLUHM LEGAL CLINIC: CTR. ON WRONGFUL CONVICTIONS, https://www.law.northwestern.edu/legalclinic/wrongfulconvictions/exonerations/patrick-pursley.html (last visited Mar. 22, 2026).

47. *See Anthony Ray Hilton*, EQUAL JUST. INITIATIVE, https://eji.org/cases/anthony-ray-hinton (last visited June 23, 2025).

48. Rehman Tungekar, *Could They Really Match Those Bullets in the Tardy Furniture Case?*, AM. PUB. MEDIA REPS. (May 8, 2018), https://www.apmreports.org/story/2018/05/08/ballistics-match-bullets-tardy-furniture.

49. *The Foundations of Firearm and Toolmark Identification,* SCI. WORKING GROUP FOR FIREARMS & TOOLMARKS, 3 (May 1, 2013), https://www.nist.gov/system/files/documents/2016/11/28/swggun_foundational_report.pdf ("With respect to tools and toolmarks, when the surface of a harder object (the tool) comes into contact with a softer object (the work piece), the harder object will impart its unique marks or features on the softer object thereby enabling a trained examiner to identify the source of a toolmark by comparing known toolmarks produced by that tool."). However, as detailed later in this Article, the foundational assumption of toolmark "uniqueness," both in firearms examination and other toolmark disciplines, remains empirically unproven.

50. A cartridge case houses the bullet until fired. When a gun is fired, the bullet is projected in the direction the gun is aimed, while the used (or "spent") cartridge casing is ejected from the firearm and typically lands on

"ballistic evidence"). Firearm experts compare ballistic evidence and attempt to draw conclusions to aid criminal investigations and prosecutions. This Subpart explains the terms necessary to understand the process of firearms examination.

Toolmarks are created every time a gun is fired. When the trigger is pulled, gas pressure is exerted in all directions, the firing pin strikes the cartridge containing the bullet, and the bullet is forced through the barrel.[51] The internal components of the gun leave behind toolmarks on the fired (bullet) and ejected (cartridge case) projectiles. Common categories of toolmarks relied upon by firearm examiners are "striations,"[52] "firing pin impressions,"[53] breech face marks," "ejector marks," and "chamber marks."[54] In plain speak, they are scratches and dents. Variation in the appearances of these scratches and dents can result from differing manufacturing designs or firearm wear and tear.[55]

Image 1. Discharged bullet (left) and discharged cartridge case and bullet (right). Striations on the bullet's surface and the firing pin impression on the casing are visible to the naked eye.[56]

Firearm examiners are overwhelmingly law enforcement officers with specialized training. They use ""comparison microscopes" that allow side-by-side microscopic comparison of evidence items.[57] The process of firearms examination generally follows one of three general scenarios. First scenario: An examiner receives a bullet or casing (projectile) from a case under investigation. The detective inspects the projectile under a microscope and draws conclusions, based on markings, about what kind of gun fired the projectile. Second scenario:

---

the ground closer to the shooter than the target. Not all guns eject cartridge cases when fired (for example, rifles), but most semi-automatic handguns do.

51. *See* BALLISTIC IMAGING, *supra* note 40, at 40–41 (providing a more complete explanation of the firing process).

52. Striations look like fine lines or scratches. *Id.* at 45 ("striations" sometimes pluralized as "striae").

53. Firing pin impressions look like circular indentations. *See id.* at 43.

54. *See id.* at 44–45; *id.*, at 41–46 (providing a more comprehensive outline of the categories and sources of firearm toolmarks on bullets and shell casings).

55. *See id.* at 74.

56. *How Good a Match is It? Putting Statistics into Forensic Firearms Identification*, NAT'L INST. OF STANDARDS & TECH. (Feb. 3, 2025), https://www.nist.gov/news-events/news/2018/02/how-good-match-it-putting-statistics-forensic-firearms-identification.

57. *See Comparison Microscope*, NAT'L INST. OF STANDARDS & TECH. (Jan. 15, 2025), https://www.nist.gov/glossary-term/37626.

An examiner receives multiple projectiles during a criminal investigation, inspects and compares the projectiles under a microscope, and concludes whether they were likely to have originated from the same gun or different guns. Third scenario: In the most common scenario, illustrated in *Ghigliotty*, an examiner receives a projectile collected in an investigation and a gun suspected to be the source. The examiner fires the suspect gun in a laboratory setting, takes the resulting "test fired" projectile, and then proceeds with the process in the second scenario in an attempt to determine whether the projectile gathered in the investigation was fired from the suspect gun. In all three scenarios, examiners use microscopes to compare toolmarks on bullets or cartridge casings.[58] Following this process, the firearm examiner authors a report with their conclusions and often testifies at trial as an expert witness.

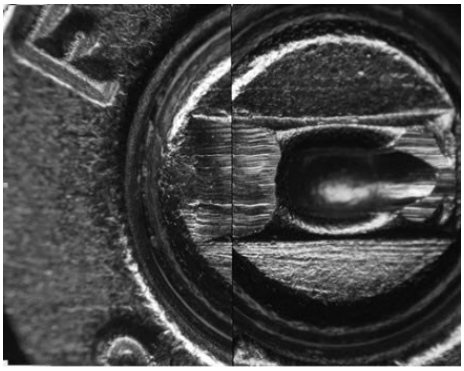


Image 2. Side-by-side microscopic comparison of two cartridge casings; the thin vertical black line divides the two casings. The casings are oriented to compare toolmarks (breech face marks and firing pin impression).[59] It is easy to see superficial similarities, which are not necessarily indicative of a shared source.

Firearms examination is part of a broader group of forensics known as "pattern-matching" or "feature-comparison," procedures though which an examiner seeks to determine whether one evidence item is associated with another based on similarity among features.[60] Other flawed pattern-matching disciplines include bite mark comparison,[61] shoeprint and tire track

58. *See* BALLISTIC IMAGING, *supra* note 40, at 65.

59. Stephen G. Bunch, Erich D. Smith, Brandon N. Giroux & Douglas P. Murphy, *Is a Match Really a Match? A Primer on the Procedures and Validity of Firearm and Toolmark Identification*, 11 FORENSIC SCI. COMM., July 2009, at 7 fig. 1.

60. *See* PCAST REPORT, *supra* note 34, at 46.

61. *See* Erica Beecher-Monas, *Reality Bites: The Illusion of Science in Bite-Mark Evidence*, 30 CARDOZO L. REV. 1369, 1372 (2009) (explaining that bitemark analysis is "a field replete with the trappings, if not the substance, of science," and describing how these "trappings of science" persuade nonscientist lawyers, judges, and juries).

comparison,[62] and fingerprint identification.[63] Like other pattern-matching experts, what firearm examiners do is actually quite straightforward. They look at two items of ballistic evidence under a microscope and decide whether they think the marks on one evidence item look the same, similar to, or different from marks on the other evidence item. There are no meaningful rules, no standards, and no way to know whether they are right or wrong.[64] One might compare the methodology to a children's matching game—"does this set of toolmarks look like that set of toolmarks?"—albeit one without a known correct answer.

Firearm examiners describe their method differently. While examiners concede they are making subjective judgments,[65] they use scientific-sounding language to describe the process. The Association of Firearm Toolmark Examiners ("AFTE") is the preeminent organization of firearm examiners.[66] Rather than establishing a testable method, AFTE publishes its *Theory of Identification as it Relates to Toolmarks*,[67] which allows examiners to conclude two projectiles share a "common origin" (meaning they were fired by or ejected from the same gun) when their toolmarks are in "sufficient agreement."[68] The theory defines "sufficient agreement" as existing "between two toolmarks [when] the agreement of individual characteristics is of a quantity and quality that the likelihood another tool could have made the mark is so remote as to be considered a practical impossibility."[69] An opinion can be drawn based merely on comparison of "two or more sets of surface contour patterns comprised of individual peaks, ridges and furrows."[70]

This theory—notably not an actual *method* with procedures, steps, or protocols—is troubling on a number of levels. The ability to discern projectiles' "common origin" is itself unproven.[71] The standard for "sufficient agreement" is self-defining and lacks any external, testable metric. Identification based on two sets alone is unexacting. And it is entirely "subjective."[72]

62. *See* NRC REPORT, *supra* note 25, at 145–50.

63. *See* WILLIAM THOMPSON, JOHN BLACK, ANIL JAIN & JOSEPH KADANE, FORENSIC SCIENCE ASSESSMENTS: A QUALITY GAP ANALYSIS, LATENT FINGERPRINT EXAMINATION, AM. ASS'N FOR THE ADVANCEMENT OF SCI. 13 (2017), http://www.aaas.org/resources/latent-fingerprint-examination.

64. *See* PCAST REPORT, *supra* note 34.

65. "[I]nterpretation of individualization/identification is subjective in nature, founded on scientific principles and based on the examiner's training and experience." *AFTE Theory of Identification as It Relates to Toolmarks*, ASS'N OF FIREARM & TOOLMARK EXAM'RS, https://temp.afte.org/about-us/what-is-afte/afte-theory-of-identification (last visited July 27, 2025).

66. *What is AFTE?*, ASS'N OF FIREARM & TOOLMARK EXAM'RS, https://afte.org/about-afte/what-is-afte/ (last visited July 24, 2025).

67. ASS'N OF FIREARM & TOOLMARK EXAM'RS, *supra* note 65.

68. *Id.*

69. *Id.*

70. *Id.*

71. *See* BALLISTIC IMAGING, *supra* note 40, at 81.

72. ASS'N OF FIREARM & TOOLMARK EXAM'RS, *supra* note 65.

The scientific principles that AFTE's theory purports to be founded on are neither articulated nor readily apparent, and scientists and experts have revealed this theory to lack scientific foundation.

2. Scientific Critique of Firearms Examination

When scientific communities took a closer look at forensic firearms examinations, critical foundational flaws were immediately apparent. In 2008, the National Research Council ("NRC") engaged in a comprehensive assessment of firearms examination.[73] While the NRC's initial goal was to explore ways to *expand* the discipline,[74] its inquiry instead revealed a lack of empirical support for the basic premises of firearms examination: uniqueness and reproducibility.[75]

The premise of "uniqueness" assumes that every firearm creates unique toolmarks, so that toolmarks created on projectiles discharged from one firearm will be distinct from toolmarks created on projectiles discharged from every other firearm in the world.[76] "Reproducibility" assumes that a firearms examination method can be reliably repeated by different examiners to reach the same result.[77] Firearm toolmark identifications can only be reliably made if firearm toolmarks are, in fact, unique and the methods of examination reproducible. But the NRC made a "clear and unambiguous" finding that the "validity of the fundamental assumptions of uniqueness and reproducibility of firearms-related toolmarks has not yet been fully demonstrated."[78] This means no one can say, regardless of methodology, whether it is even *possible* to correctly match bullets and cartridge casings to the gun that fired them. This finding was supported and echoed by subsequent scientific inquiries.

In another report authored a year later, the NRC wrote "[s]ufficient studies have not been done to understand the reliability and repeatability of the methods" of firearms examination.[79] The NRC report also noted "the lack of a precisely defined process . . . [that] does not even consider, let alone address, questions regarding variability, reliability, repeatability, or the number of correlations needed to achieve a given degree of confidence."[80]

---

73. *See* BALLISTIC IMAGING, *supra* note 40.

74. *Id.* at 16–17 (outlining the policy objectives guiding the NRC's inquiry, including maintenance and enhancement of the existing firearm database ("NIBIN"), and establishing a "national reference ballistic image database . . . containing images of ballistic samples from all newly manufactured or imported guns . . . to generate investigative leads from the point of sale of a firearm.").

75. *Id.* at 3.

76. The assumption of toolmark uniqueness has been described as "the principle of uniqueness . . . wherein, all objects are unique to themselves and thus can be differentiated from one another." PCAST REPORT, *supra* note 34, at 61 n.149.

77. *See* PCAST REPORT, *supra* note 34, at 47 ("By 'reproducible,' we mean that, with known probability, different examiners obtain the same result, when analyzing the same samples.").

78. *See* BALLISTIC IMAGING, *supra* note 40, at 81.

79. *See* NRC REPORT, *supra* note 25, at 154.

80. *Id.* at 155.

Scientific critique continued. In 2016, the President's Council of Advisors on Science and Technology ("PCAST")[81] issued a report excoriating nearly every forensic discipline used in criminal investigations and prosecutions, including firearms toolmark examination. Citing many of the concerns raised by the NRC, the PCAST report identified and examined the scientific criteria required for foundational validity:

> (1) [A] reproducible and consistent procedure for (a) identifying features . . . ; (b) comparing the features . . . ; and (c) determining . . . whether the samples should be declared to be a proposed identification . . . [and] (2) empirical measurements, from multiple independent studies, of (a) the method's false positive rate—that is, the probability it declares a proposed identification between samples that actually come from different sources and (b) the method's sensitivity—that is, probability that it declares a proposed identification between samples that actually come from the same source.[82]

PCAST considered each criterion in turn, finding the discipline deficient on both. Looking first to AFTE's theory as the only "procedure" available, they noted its troubling circularity, immunizing conclusion from any meaningful external metric,[83] its entirely subjective nature, and the lack of any protections from vulnerability to human error, inconsistency, and cognitive bias.[84] Turning to the second criterion, empirical measurements, PCAST found only *one* error rate study appropriately designed to test foundational validity and estimate

81. The President's Council of Advisors on Science and Technology comprises representatives from diverse fields, including academic researchers from major universities (Harvard, MIT, Princeton, University of California system, Northwestern, and others) spanning disciplines such as physics, chemistry, biology, computer science, engineering, and environmental science; industry executives from aerospace, biotechnology, venture capital, and technology sectors; healthcare and medical institution leaders; and individuals with government science policy expertise. Senior advisors, including federal judges, law school deans and professors, and expert statisticians, supported the council. *See* PCAST REPORT, *supra* note 34, at v–ix.

82. *Id.* at 48.

83. *See id.* at 104 ("The 'theory' states that an examiner may conclude that two items have a common origin if their marks are in 'sufficient agreement,' where 'sufficient agreement' is defined as the examiner being convinced that the items are extremely unlikely to have a different origin.").

84. *Id.* at 49. Cognitive bias may include contextual bias, confirmation bias, and avoidance of cognitive dissonance. "Contextual bias . . . occurs when decisionmakers are influenced by exposure to extraneous information that is not necessary to make the decision at hand." Elizabeth J. Reese, Comment, *Techniques for Mitigating Cognitive Biases in Fingerprint Identification*, 59 UCLA L. REV. 1252, 1260 (2012). For decades, cognitive psychologists have identified contextual bias as a source of error in human decision-making. *See id.* at 1258–61. Confirmation bias exists where "individuals interpret information, or look for new evidence, in a way that conforms to their pre-existing beliefs or assumptions." PCAST REPORT, *supra* note 34, at 31. Avoidance of cognitive dissonance is an individual's reluctance to "accept new information that is inconsistent with their tentative conclusions. *Id.*; Itiel E. Dror, David Charlton & Ailsa E. Péron, *Contextual Information Renders Experts Vulnerable to Making Erroneous Identifications*, 156 FORENSIC SCI. INT'L 74, 74–78 (2006). Errors that result from exposure to task-irrelevant, biasing information do not reflect ill-intent on the part of the decision-maker, but it is the very unconscious nature of these biasing effects that make them so pernicious in the courtroom. As researchers Dror and Cole stated, "[e]ven more than an honestly mistaken eyewitness, an honestly mistaken expert is the least culpable and thus, potentially, the most dangerous kind of witness that can testify in a legal proceeding." Itiel E. Dror & Simon A. Cole, *The Vision in "Blind" Justice: Expert Perception, Judgment, and Visual Cognition in Forensic Pattern Recognition*, 17 PSYCHONOMIC BULL. & REV. 161, 162 (2010).

reliability, falling far short of establishing foundational validity.[85] Even that one study had critical design flaws.[86] In recent years, researchers and statisticians have more closely examined the research designs of firearms examination error rate studies, providing further insight into their methodological shortcomings.[87] A 2024 comprehensive analysis of existing error rate studies concluded that *all* of these studies are so gravely methodologically flawed that their results are rendered scientifically invalid, and "incapable of establishing scientifically validity of the field of firearms examination."[88]

The PCAST report echoed NRC findings, concluding that "firearms analysis currently falls short of the criteria for foundational validity."[89] Subsequent scientific inquiry has revealed it still falls short.[90] Notably, while the NRC and PCAST called for more rigorous empirical testing, there is reason to question whether firearm toolmark analysis can *ever* be validated. While the assumption of uniqueness is empirically unproven (there is insufficient data available to demonstrate its truth), the assumption of reproducibility is directly refuted by the body of knowledge about firearms examination. The same firearm may—and often does— create different toolmarks on different projectiles based on a number of factors, primarily degradation of the firearm's barrel due to the passage of time, manner of storage, use, humidity, or cleaning.[91] Accordingly, not only is the validity of firearms examination yet unproven, there is reason to question the ability to establish foundational validity of a field that purports to individualize projectile toolmarks to specific firearms.

85. *See* PCAST REPORT, *supra* note 34, at 111.

86. Even the "Ames Laboratory study" (finding an error rate of 1 in 66 but discounting 33.7% due to "inconclusive" responses), is subject to shortcomings characteristic of forensic error rate testing. *Id.*, at 111, tbl. 2.

87. *See generally* Maria Cuellar, Susan Vanderplas, Amanda Luby & Michael Rosenblum, *Methodological Problems In Every Black-Box Study of Forensic Firearm Comparisons*, 23 L., PROBABILITY & RISK, Dec. 2024, at 1, 1–23 (2024) (concluding that every firearms examination error rate study reviewed in the literature suffered from such fatal methodological flaws that their conclusions are invalid and fail to validate the procedures of firearms examination); *see also* Kori Khan & Alicia Carriquiry, *Shining a Light on Forensic Black-Box Studies*, 10 STAT. & PUB. POL'Y, 2023, at 1 (discussing methodological flaws in firearms examination error rate studies).

88. *See id.* at 2 tbl. 1 (identifying six distinct flaws observed in error rate studies—inadequate sample size, non-representative samples, non-representative testing conditions and environment, inconclusive responses are treated as correct or ignored, invalid or nonexistent uncertainty measures for error rates, and missing data—and the impact of these flaws.)

89. *See* PCAST REPORT, *supra* note 33, at 112.

90. *See* Cuellar *supra* note 87; *supra* note 88.

91. The assumptions of reproducibility in firearms examination are contradicted by research, showing that factors such as barrel wear, storage conditions, and repeated use cause firearms to produce varying marking over time, undermining the reliability of ballistic identification. *See* BALLISTIC IMAGING, *supra* note 40, at 74. Thus, marks left by the same firearm can change significantly over its operational lifetime, making it increasingly difficult to establish definitive matches as wear progresses. *See* Bunch et al., *supra* note 59, at 10–11.

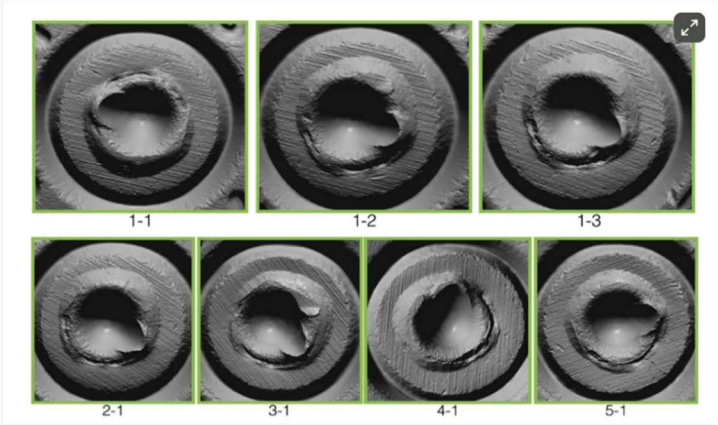


Image 3. Images of seven ejected casings fired from the same gun, illustrating how the same firearm can create toolmarks that, even to the untrained eye, clearly differ.[92]

After over a century as a fixture in criminal investigations, scientific inquiry revealed firearms examination as junk science. Then came automation.

B. VIRTUAL COMPARISON MICROSCOPY: AUTOMATING JUNK SCIENCE

Firearms examinations are increasingly automated by algorithms that create, modify, and compare digital 3D images of ballistic evidence. While developers offer a suite of functions, this Article focuses primarily on 3D scanning and virtual comparison microscopy (VCM) systems.[93] VCM has, for the most part, crept quietly into the landscape of automated criminal forensic tools, leaving the spotlight on more well-known policing technologies such as algorithmic risk assessment tools[94] and facial recognition technology.[95] The

92. Balko, *supra* note 38; Pierre Duez, Todd Weller, Marcus Brubaker, Richard E. Hockensmith & Ryan Lilien, *Development and Validation of a Virtual Examination Tool for Firearm Forensics*, 63 J. FORENSIC SCI. 1069 (2018).

93. The scope of automated firearms examination technology is broader than discussed in this Article. Functions include, at least, 3D imaging/scanning systems, "triage," "database search," data sharing systems, report generation functions. This Article uses the term "VCM" broadly, while acknowledging this does not capture the breadth of the phenomenon. *See* TopMatch-3D: High Capacity, CADRE FORENSICS, https://www.cadreforensics.com/pdf/TopMatch-3D-HighCapacity_June2022.pdf (last visited Jan. 5, 2026).

94. *See* Richard A. Webster, *An Algorithm Deemed This Nearly Blind 70-Year-Old Prisoner a "Moderate Risk." Now He's No Longer Eligible for Parole.*, PROPUBLICA (Apr. 10, 2025, at 06:00 ET), https://www.propublica.org/article/tiger-algorithm-louisiana-parole-calvin-alexander (reporting on Louisiana's use of algorithms such as the Targeted Interventions to Greater Enhance Re-entry ("TIGER") in parole determinations as part of Governor Jeff Landry's tough-on-crime agenda, making Louisiana the only state to use risk scores to automatically rule out large portions of prisoners from parole consideration, despite the tool originally being designed for rehabilitation rather than as a punitive measure); Chelsea Barabas, Karthik Dinakar & Colin Doyle, *The Problems with Risk Assessment Tools*, N.Y. TIMES (July 17, 2019), https://www.nytimes.com/2019/07/17/opinion/pretrial-ai.html.

95. *See generally* ALL THINGS CONSIDERED: *The Debate Over Facial Recognition Technology's Role in Law Enforcement*, NPR (July 10, 2019, at 17:48 ET), https://www.npr.org/2019/07/10/740480966/the-debate-

ubiquitousness of VCM in criminal investigations, however, places on the legal community—litigants, courts, legislators, and academics alike—an imperative to pull back the curtain and become familiar with the nature of algorithmic technology and its use, risks, and impact.

1. Virtual Comparison Microscopy 101

BULLETTRAX, used in *Ghigliotty,* is one of many automated tools transforming firearms examination by digitizing, modifying, and examining ballistic evidence through algorithmic function. This Subpart provides a broad overview of how emerging forensic technologies—VCM and 3D scanners—work. These tools offer distinct functions, all enabled by algorithms and can be used alone or in tandem.[96]

First, 3D scanners create digital images of firearms projectile evidence. BULLETTRAX is an example of a 3D scanner. Unlike microscopes that magnify evidence items, 3D scanners create digital images of evidence, which may then be modified and examined with companion VCM functions.

Virtual 3D images created by machines like BULLETTRAX are not photographs of the evidence; they are computer-generated images created from data inputs, algorithmic interpretation, and human modification. These machines use computer software to create and modify computer-generated images of the surface of a bullet or cartridge casing. Virtual imaging tools "flatten" the cylindrical surface of a bullet so that an examiner can look at the entire surface at the same time. Forensic companies developing these tools claim they "capture[] ultra-fine bullet details at the submicron level with exceptional clarity and precision,"[97] "measure accurate 3D surface topographies in standard units

over-facial-recognition-technologys-role-in-law-enforcement (debating the use of facial recognition technology); *e.g.*, Andrea May Sahouri, *Lawsuit Filed After Facial Recognition Tech Causes Wrongful Arrest of Pregnant Woman*, USA TODAY (Aug. 8, 2023, at 13:49 ET), https://www.usatoday.com/story/news/nation/2023/08/08/facial-recognition-technology-wrongful-arrest-pregnant-woman/70551497007 (reporting on the wrongful arrest of an eight-month pregnant Black woman in Detroit, falsely identified by facial recognition technology and arrested for carjacking and robbery, marking the sixth documented case of false accusations due to facial recognition technology, with all six victims being Black individuals, and the third such case in Detroit); Sophie Wentzell, *ACLU Criticizes New Orleans Police for Using Facial Recognition in Secret*, VANGUARD NEWS GRP. (May 22, 2025), https://davisvanguard.org/2025/05/aclu-accuses-nopd-facial-recognition (reporting on the ACLU's call for the New Orleans Police Department to cease using facial recognition technology through Project NOLA's surveillance camera network, which has installed over 200 facial recognition-enabled cameras throughout the city since 2023).

96. Some, such as the Integrated Ballistic Identification System ("IBIS"), separate functions into different machines, using dedicated acquisition stations for capturing images, correlation engines for running comparisons, and distinct review workstations for examining results. *See Firearm & Tool Mark Identification–IBIS*, LEADSONLINE, https://leadsonline.com/ibis (last visited Jan. 7, 2026). Others offer both functions in the same machine, such as the TopMatch-3D High Capacity system which integrates 3D scanning, VCM, database search, automated comparison, and analysis tools into a single computer workstation. *See TopMatch-3D: High Capacity*, CADRE FORENSICS, https://www.cadreforensics.com/pdf/TopMatch-3D-HighCapacity_June2022.pdf (last visited Jan. 7, 2026).

97. *See BulletTrax*, LEADSONLINE, https://leadsonline.com/bullettrax (last visited Jan. 7, 2026).

resulting in a detailed heightmap of the cartridge case surface,"[98] "offer examiners significantly more detail than traditional 2D images,"[99] and "employ[] advanced surface tracking technology to accurately adapt to any bullet's contours, regardless of condition."[100]

Different virtual imaging tools use different methods[101] to gather topographical data from the surface of a projectile, but all are controlled by proprietary algorithms.[102] Forensic technology developers boast advantages over traditional methods in addressing common comparison challenges, such as damaged projectiles, lighting conditions, steep slopes, or the reflective nature of metal.[103] Each of these "advantages" necessarily alters the resulting image.

Second, VCM software allows operators to use "filters" to modify the virtual images. Common filters adjust lighting conditions and remove a bullet's round shape, "noise,"[104] and irregularities from a bullet's surface.[105] These algorithm-enabled filters visually alter the image created of the projectile's surface.

98. RYAN LILIEN, FIREARM FORENSICS BLACK-BOX STUDIES FOR EXAMINERS AND ALGORITHMS USING MEASURED 3D SURFACE TOPOGRAPHIES 2 (2019), https://www.ojp.gov/pdffiles1/nij/grants/254338.pdf.

99. *Id.*

100. LEADSONLINE, *supra* note 97.

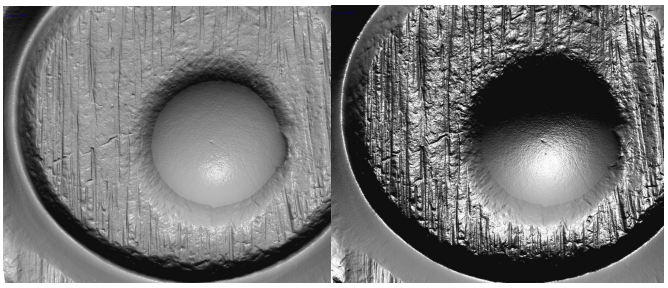
101. Different methods include confocality, interferometry, and focus variation. *See* T. Brian Renegar, *Implementation of 3D Technology, Analysis, and Statistics for Firearm and Tool Mark Examinations: 3D Instrument Measurement*, NIST FORENSIC SCIS. PRESENTATION (May 26, 2019) [hereinafter *Implementation of 3D Technology*], (PowerPoint slides provided to attendees of the 2019 AFTE Annual Conference, on file with the author); *Focus Variation Microscope*, NAT'L INST. OF STANDARDS & TECH. (Mar. 7, 2025), https://www.nist.gov/laboratories/tools-instruments/focus-variation-microscope; *see also Focus Variation*, BRUKER ALICONA, https://www.alicona.com/en/technologies/focus-variation (last visited Jan. 7, 2026).

102. *See Cadre's 3D Scan Acquisition Technology*, CADRE FORENSICS, https://www.cadreforensics.com/technology.html (last visited Jan. 7, 2026) (explaining that Cadre Forensics' 3D Scan Acquisition Technology also includes a proprietary gel pad, which "temporarily deforms to the shape of any object pressed into its surface. . . . remov[ing] surface specularities from any material (including metal)").

103. *Implementation of 3D Technology*, *supra* note 101.

104. *See* Prabu Kumar, *A Deep Dive into Types of Camera Noise, and Their Impact on Image Quality*, E-CON SYS. (Apr. 23, 2024), https://www.e-consystems.com/blog/camera/technology/a-deep-dive-into-types-of-camera-noise-and-their-impact-on-image-quality ("[N]oise refers to any unwanted signal variation in captured images that does not originate from the scene being imaged. This noise can manifest as random pixel variations (graininess) or color artifacts, detracting from the accuracy of the image data and lowering the system's effectiveness.").

105. *Implementation of 3D Technology*, *supra* note 101.



Images 4 and 5. VCM-created images from Cadre Forensics website[106] depict two views of the same computer-generated rendering of a fired cartridge casing, with differing lighting settings. The image on the left is in normal mode. The image on the right is the same rendering in "enhanced contrast mode."

VCM also allows users to modify the presentation of images. They can "load scans side-by-side, rotate and translate the scans in locked or unlocked mode, zoom into fine detail, adjust lighting, annotate surface detail, and export high-resolution screenshots for inclusion into reports and presentation."[107] Some tools offer color coding of topography areas of similarity to "aid"—or guide "like a GPS" as in *Ghigliotty*[108]—examiners' comparisons.[109] These options lead to modern, visually appealing presentations that can provide an unfounded sense of legitimacy.

Third, the ultimate function of VCM is automated evidence comparison: The machine itself compares images of evidence items and concludes that their toolmarks are so similar that (in the machine's "opinion") they were fired by the same gun, that they are so dissimilar they could not have been fired by the same gun, or somewhere in between—"inconclusive."[110]

106. *See Cadre-VCM: Validated Virtual Comparison Microscopy (VCM)*, CADRE FORENSICS, https://www.cadreforensics.com/VirtualComparisonMicroscopy.html (last visited Jan. 7, 2026).

107. *Id.*

108. *See, e.g.*, State v. Ghigliotty, 232 A.3d 468, 477 (N.J. Super. Ct. App. Div. 2020) (Sandford used the screenshots from BULLETTRAX "like a GPS" to guide his comparison microscope analysis of the areas of interest (citation modified)).

109. *See Cadre's Suite of Validated 3D Imaging and Analysis Software for VCM*, CADRE FORENSICS, https://www.cadreforensics.com/Software.html (last visited Jan. 7, 2026).

110. *See* RONALD NICHOLS, BUILDING A PREVENTIVE CRIME GUN STRATEGY: A PLAYBOOK FOR SUCCESS 14–16 (2d ed. 2020); *Enhancing Gun Crime Investigations Through Automated Ballistic Identification*, LEADSONLINE 3–4, https://46525538.fs1.hubspotusercontent-na1.net/hubfs/46525538/FTLO_WhitePaper_ok.pdf (last visited Jan. 7, 2026). These automated systems generate similarity scores—for example, "same source," "different source," or "inconclusive"—between toolmarks using proprietary algorithms whose internal mechanics are unknown to the scientific or legal community.

2. The Role of Virtual Comparison Microscopy in Criminal Prosecutions

Virtual comparison microscopy can be used by law enforcement in a number of ways. It is most commonly used to create and compare virtual images with ballistic databases to develop investigative leads.[111] It can also be used, as in *Ghigliotty*, to create and manipulate virtual images used in a human expert's examination. VCM can be used as a "blind verification tool" to check a human expert's conclusion.[112] Finally, VCM can be used to fully automate the comparison of ballistic evidence. While it is impossible to gather complete information regarding the pervasiveness of VCM in policing and prosecutions,[113] it is apparent that the technology has long been used in the shadows[114] of criminal investigations and, more recently, as a more central forensic examination tool. Now proponents are working toward expanding its role in criminal courtrooms.[115] The implications of using VCM to automate flawed firearms examination methods are illustrative of the broader consequences of automating forensics.

II. HOW AUTOMATION CHANGES (AND DOES NOT CHANGE) JUNK SCIENCE

Automation changes the process of examining and creating forensic evidence; it does not change the nature of the evidence or underlying methods. Therefore, while offered as a response to the scientific deficits of faulty forensic methods, algorithms have proven to introduce more problems than solutions. Through an analysis of VCM's impact on firearm examination, this Part reviews the ways in which automation risks obscuring the deficiencies of junk science and retains underlying issues with subjectivity and foundational invalidity, while also adding opportunities for computer code, machine, or operator error and bias undermining its conclusions.

111. *See infra* Subpart.III.B.3.

112. Xiaoyu Alan Zheng & Doug Lancon, *Analysis of Breech Face Marks with Heavy Subclass Influence Through Traditional and Novel Methods*, AFTE (2024), https://afte.org/shop/training-seminar-media/training-seminar-usb-afte-2024-anchorage-ak (presentation at the AFTE Seminar in Anchorage, Alaska) (video on file with the author).

113. Among other reasons for this absence of information, the use of investigative tools is often not disclosed to defense attorneys and courts and most trial court decisions are unpublished.

114. *See* Lau, *supra* note 18 (manuscript at 41) ("Because FRE 702 and *Daubert* do not govern the admissibility of evidence at pretrial hearings, or at the investigative stage, and because 99% of criminal cases do not go to trial, the use of call analysis [or other forensic practices] may go undetected, and therefore, unscrutinized in in the vast majority of cases where it is deployed.").

115. *See A Century of Ballistics Comparison Giving Way to Virtual 3D Methods*, NAT'L INST. OF JUST. (Mar. 23, 2022), https://nij.ojp.gov/topics/articles/century-ballistics-comparison-giving-way-virtual-3d-methods ("Zheng, Soons, and Lilien all agree that it could be three to five years before the results of the 3D scans are routinely accepted by the courts.").

A. FAÇADE OF TECH LEGITIMACY OBSCURES SCIENTIFIC DEFICITS

Automation does not *un-junk* junk science. Rather, it risks obscuring flaws and threatens to undo decades of work uncovering scientific deficits in disciplines masquerading as science. The impact of automation on the perception of junk science, erecting a "façade of tech legitimacy," is clearly illustrated in *Ghigliotty*. Two firearms examinations were conducted by trained experts. One employed traditional, human-led and microscope-aided methods. The second relied on automation. Lacking any information about how (or whether) the technology worked, law enforcement, prosecutors, and judges all assumed that the latter conclusion was the more legitimate and correct conclusion simply due to its use of novel, algorithmic technology.[116] This is illustrative of the broader phenomenon of automation bias.[117] Humans tend to perceive machines and algorithms as objective, reliable, and not subject to the whims, biases, and missteps of human error.[118] These perceptions are misguided, as computers are subject to the whims, biases, and missteps of both coders and users,[119] as well as computer and machine malfunctions.[120] Yet they endure.

Capitalizing on automation bias threatens to erase progress made toward exposing deficiencies in forensic methods. Just as skepticism around traditional method of firearms examination was finally gaining hold,[121] novel technology was advanced as a solution. Instead, it offers a shroud of false validity, repeating and reanimating a dangerous history. The façade of tech legitimacy echoes a façade of scientific legitimacy that contributed to junk science's rise to prominence in criminal courts. Throughout the meteoric expansion of forensic "sciences" in criminal investigations and prosecutions in the twentieth century, we repeatedly saw the human tendency to defer to "science"—or, what *sounded*

116. This was evident from the manner with which the parties discussed evidence in the case, but it was most clearly apparent simply from the facts of these parties' decisions to arrest, charge, indict, and set an unattainable bail for Mr. Ghigliotty based entirely on the BULLETTRAX-aided expert conclusion that directly contradicted the prior non-automated expert conclusion.

117. *See* Lauren Kahn, Emelia S. Probasco & Ronnie Kinoshita, *AI Safety and Automation Bias*, CTR. SEC. & EMERGING TECH., Nov. 2024, at 1, https://cset.georgetown.edu/wp-content/uploads/CSET-AI-Safety-and-Automation-Bias.pdf (defining "automation bias" as the "tendency for an individual to over-rely on an automated system."); *see also* Bryce Hoffman, *Automation Bias: What It Is and How to Overcome It*, FORBES (Mar. 10, 2024, at 1:52 ET), https://www.forbes.com/sites/brycehoffman/2024/03/10/automation-bias-what-it-is-and-how-to-overcome-it.

118. This deference to automated machines is amplified as subject matter becomes more complex. *See generally* Eric Bogert, Aaron Schecter & Richard T. Watson, *Humans Rely More on Algorithms Than Social Influence as a Task Becomes More Difficult*, 11 SCI. REP. 8028 (2021), https://doi.org/10.1038/s41598-021-87480-9 (reporting research from experiments finding human participants relied more on algorithmic advice than social influence, and increasingly so as tasks became more difficult).

119. *See infra* Subpart.II.C, E.

120. *See infra* Subpart.II.E.

121. *See* Garrett et al., *supra* note 42, at 128–40 (discussing post-*Daubert* court scrutiny of firearm toolmark evidence).

like science despite lacking any meritorious features of true science.[122] No doubt aided by the mass popularity of television shows like CSI,[123] forensic evidence is often the most persuasive evidence to jurors.[124] The rise of DNA exonerations reveals in striking detail the danger of such blind deference, while the NRC and PCAST reports explain that forensic "sciences" are overwhelmingly not science at all.[125]

Uncritical reliance on automated forensic tools risks repeating this history. Coining complex terminology and cloaking methodology with opacity only increases the risk that prosecutors will simply *assume* objectivity and accuracy and defer to those assumptions in charging decisions. That risk carries through to defense attorneys, who may be dissuaded from mounting legal challenges, judges who may inadequately scrutinize the evidence, and jurors who may defer to computers in rendering verdicts. The extent to which automation discourages rigorous testing, open review, and candid discussion of its limitations is a monumental concern that pervades forensic evidence broadly.

When one looks beyond the façade of legitimacy, it becomes clear that automation raises numerous issues with legal and scientific implications, some artifacts of the underlying "science," and many novel features of automation.

B. AUTOMATION DOES NOT ELIMINATE SUBJECTIVITY

Despite claims to the contrary,[126] automation does not remove subjectivity from forensic analyses. A central critique of traditional firearms examination theory is its subjective nature, and, relatedly, the substantial threat of cognitive bias skewing conclusions.[127] Replacing a human with a machine necessarily controls *some* degree[128] of subjectivity and bias at *one* inflection point.[129] Yet this single point of the examination is colored by countless others, all replete with subjective choices.

---

122. Sinha, *supra* note 25, at 882 ("Ironically, the extent to which lay people without scientific training tend to trust forensic science evidence and mistakenly believe that it brings neutrality, fairness, accuracy, and certainty to the criminal process has allowed forensic evidence to do just the opposite.").

123. J. Herbie DiFonzo & Ruth C. Stern, *Devil in a White Coat: The Temptation of Forensic Evidence in the Age of CSI*, 41 NEW ENG. L. REV. 503, 505–06 (2007).

124. *See* Sinha, *supra* note 25, at 881.

125. *See generally* NRC REPORT, *supra* note 25; PCAST REPORT, *supra* note 33. Subsequent researchers and statisticians have continued to expose the ways that forensic methods routinely accepted in criminal courts are subject to critical foundational flaws. *See, e.g.,* Cuellar et al., *supra* note 84 (discussing the design flaws that undermine the results of firearms examination validation studies that have been routinely relied upon to claim the discipline's validity).

126. *Implementation of 3D Technology, supra* note 103, at slide 10 (listing "facilitates objective comparisons" as a "benefit" of 3D firearm imaging technology, as compared with traditional "2D" methods).

127. *See supra* Subpart.I.A–B; PCAST REPORT, *supra* note 33.

128. As discussed in this Subpart, algorithms are vulnerable to the impact of subjectivity and bias of computer coders.

129. Instead of a human firearm expert looking for toolmarks with evidentiary significance, VCM creates an image without an explicit case agenda. Instead of a human firearm expert making a comparison of bullets with the hope of solving a case, VCM compares microscopic markings for similarity or dissimilarity, without hopes or wishes.

First, automated firearms examination technology only responds to subjectivity concerns when used for comparison. When used (as it primarily is now) for scanning and modification, human examiners remain the decision makers.[130] Even with automated comparison, humans determine machine inputs and shape machine outputs. On the front end, humans (police, detectives, forensic experts, prosecutors) make decisions about which investigation leads to pursue, who to arrest, what evidence to collect, and which evidence to submit for automated examination. As other scholars have put it, "raw data is an oxymoron"[131]: "[A]ll machine output reflects human choices about input, "just as a direct examination of a witness in a justice's parlor reflects choices about what questions to ask."[132] These subjective human decisions necessarily narrow the universe of options available for the algorithm's consideration. On the back end, human examiners make decisions about machine settings and interpret results.

Subjective human decisions also impact the *appearance* of images created by 3D scanners and VCM, which might be relied upon by examiners and entered as evidence. Human users decide which filters to apply and how to adjust settings. Their decisions may be shaped by cognitive bias or intentional manipulation, particularly if decisions are made by a firearm examiner with case-specific information.[133] Even a well-meaning firearm examiner might be more likely to apply filters that create images confirming or bolstering a desirable or expected result—for example, filters that make two items of evidence look more similar—due to the influence of confirmation bias.[134] Less well-meaning examiners might intentionally manipulate images to better reflect their desired result.

Even the conclusions of algorithms are not entirely objective. Machines can be coded to adopt different assumptions and means of interpreting results,[135] which can lead to different outputs with different legal implications.

This was illustrated in a high-profile case involving automated DNA analysis technology, where a changed reporting threshold altered the outcome of a case. Nick Hillary was arrested for the 2011 killing of his ex-girlfriend's son. DNA samples taken from Hillary's coffee cup, car, home, and clothes and compared with dozens of DNA samples from the crime scene and the decedent's

130. *See, e.g.*, State v. Ghigliotty, 232 A.3d 468, 474 (N.J. Super. Ct. App. Div. 2020) ("When Sandford returned to his lab . . . he used the screen shots from BULLETTRAX like a GPS to guide his comparison microscope analysis of the areas of interest and came to an opinion of an identification or a positive identification." (citation modified)).

131. *See* generally DANIEL ROSENBERG ET AL., "RAW DATA" IS AN OXYMORON (Lisa Gitelman ed. 2013) (collecting essays exploring how the generation and interpretation of data is culturally determined).

132. Andrea Roth, *Machine Testimony*, 126 YALE L.J. 1972, 2043 (2017).

133. Firearm examiners are routinely offered case details irrelevant to the examination. Research shows this impacts forensic outcomes. *See Reese*, supra note 84, at 1260; PCAST REPORT, *supra* note 34, at 31.

134. *See* Dror et al., *supra* note 84, at 74-78; Dror & Cole, *supra* note 84, at 162.

135. Illustrated below in the case of Nick Hillary, in which two different automated DNA analysis tools offered differing degrees of sensitivity in reporting results.

body by a human chemist through the traditional methods.[136] There were no matching results, nor any that conclusively excluded Hillary as a DNA contributor.[137] Law enforcement then sent the evidence to be tested by a probabilistic genotyping tool called "TrueAllele," which employs algorithms to automate DNA comparisons. The TrueAllele computer concluded there was "no statistical support for a match," but a year later a new prosecutor submitted the same evidence to a competitor company using technology called STRmix. STRmix reported that Hillary was 300,000 times more likely than a random person to have contributed to the DNA mixture.[138]

TrueAllele and STRmix are both machines controlled by algorithms designed to complete the same function—interpretation of DNA evidence—but they were coded with different assumptions for statistical significance and different means of reporting those results. Arguably, divorced from their import to a criminal case, the two results were not necessarily at odds. TrueAllele reported no statistical support to conclude there was a match, while STRmix calculated the statistical probability of match. Neither concluded there was a match, and neither concluded a match was impossible. But implications of the different reporting thresholds, chosen by developers and coded by computer programmers, were monumental. Machines are not inherently objective; their programming and operation retain subjectivity.

Each subjective decision made by a coder or operator provides an opportunity for error, bias,[139] inconsistency, or misinterpretation. Complete elimination of subjectivity and human control may be an impossible, and not necessarily desirable, benchmark for forensics. But if developers truly want to make forensic tools more objective, accurate, and trustworthy, they must first be

136. *See* Jesse McKinley, *Tensions Simmer as a Small Town Seeks Answers in a Boy's Killing*, N.Y. TIMES (Mar. 5, 2016), http://www.nytimes.com/2016/03/06/nyregion/murder-of-garrett-phillips-in-potsdam-new-york.html.

137. *Id.*

138. Notice of Motion to Preclude at 8–9, New York v. Hillary, No. 2015-15, (N.Y. St. Lawrence Cty. Ct., May 31, 2016). The STRmix conclusion was offered as evidence against Hillary. A trial judge excluded the STRmix results under *Frye*. Decision & Order at 10, New York v. Hillary, No. 2015-15, (N.Y. St. Lawrence Cty. Ct. Aug. 26, 2016).

139. Algorithms have been shown to adopt explicit and systemic biases, particularly racial bias, due to bias in computer coding, data input, and research design. *See generally* Beth Findley, *Why Racial Bias Is Prevalent in Facial Recognition Technology*, HARV. J.L. & TECH. (Nov. 3, 2020), https://jolt.law.harvard.edu/digest/why-racial-bias-is-prevalent-in-facial-recognition-technology (facial recognition technology has been revealed to adopt the racial biases of the images it is "trained" on); Julia Angwin, Jeff Larson, Surya Mattu & Lauren Kirchner, *Machine Bias*, PROPUBLICA (May 23, 2016), https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing (exposing racial bias in risk assessment algorithms); Filipa Queirós, *The Visibilities and Invisibilities of Race Entangled with Forensic DNA Phenotyping Technology*, 68 J. FORENSIC & LEGAL MED., Aug. 2019 (examining the ways that DNA software adopts and exacerbates racial inequities); *Racial Disparities Discovered Among Genomic Sequencing Data*, MAYO CLINIC (July 16, 2022), https://www.mayoclinic.org/medical-professionals/cancer/news/racial-disparities-discovered-among-genomic-sequencing-data/mac-20534960 ("The reference genome itself, however, has racial bias because it came from patients mostly of European ancestry."). Notably, while software has historically been created by human coders, AI is increasingly taking over the task of creating computer code. Though beyond the scope of this Article, this reality demands additional scrutiny into AI coding and how this impacts forensic methods and the rights of the accused.

candid about their limits. Instead, they routinely falsely claim objectivity,[140] contributing to the risk of automation's false legitimizing effect.

C. AUTOMATION DOES NOT ADDRESS FOUNDATIONAL VALIDITY

Automation does not address the foundational deficiencies of junk sciences like firearms examination.[141] Regardless of method—2D, 3D, human, or machine—there remains insufficient empirical support to establish the necessary assumptions of toolmark uniqueness and reproducibility.[142] While these tools may offer some insight into the discernability of features, it remains unknown whether individual firearms create unique toolmarks. Throughout the recent history of legal challenges to the reliability of firearms examinations, some judges have recognized the nonexistent methodology, lack of meaningful peer review, and subjective process.[143] Less attention has been paid to its arguably most fatal critique: We simply have no idea if anyone—skilled or unskilled, human or machine, accurate or inaccurate—is capable of "matching" projectiles to specific firearms. While algorithms change the process, they do not change the "science." Automated junk science is still junk science.

D. AUTOMATION MAY UNDERMINE RELIABILITY

Automation introduces novel issues impacting reliability: the potential for computer code, product design, machine function, and machine operation errors. Human mistakes and biases can be written into codes that conduct complicated calculations and inform forensic decision-making and conclusions. Computer coding errors can lead to calamitous results, and they are not unique to courtrooms; coding errors have been responsible for plane crashes,[144] failed rocket launches,[145] massive security breaches.[146] Forensic tools are no exception. Defense access to computer codes uncovered coding errors impacting

140. *Implementation of 3D Technology, supra* note 103, at slide 10 (listing "facilitates objective comparisons" as a "benefit" of 3D firearm imaging technology, as compared with traditional "2D" methods).

141. *See supra* Subpart.I.B (discussing findings in scientific reports that the foundational validity of firearm and toolmark examination has not been empirically demonstrated).

142. *See supra* Subpart.I.B.

143. *See* Garrett et al., *supra* note 42, at 128–140 (discussing post-*Daubert* court scrutiny of firearm toolmark evidence).

144. Two commercial airline crashes of the same model of Boeing 737 MAX occurred within months of each other, killing 346 passengers. Reports demonstrate that an error in an automated system was at fault. *See* Andy Pasztor & Robert Wall, *As Flight-Control System Is Blamed for Boeing Crash, Pilots' Actions Also Prompt Questions*, WALL ST. J. (Apr. 8, 2019, at 22:33 ET), https://www.wsj.com/articles/as-flight-control-system-is-blamed-for-boeing-crash-pilots-actions-also-prompt-questions-11554761918.

145. *See* James Gleick, *Little Bug, Big Bang*, N.Y. TIMES (Dec. 1, 1996), https://www.nytimes.com/1996/12/01/magazine/little-bug-big-bang.html (reporting that an explosion that destroyed the Ariane 5, a rocket valued at $7 billion was due to a simple source code error).

146. A program's source code errors compromised the security of up to two-thirds of all e-commerce websites for years due to the "heartbleed" bug that went unnoticed by companies as technologically savvy and powerful as Google, Yahoo, Facebook, and Amazon. *See* Nicole Perlroth, *Experts Find a Door Ajar in an Internet Security Method Thought Safe*, N.Y. TIMES (Apr. 8, 2014, at 17:08 ET) https://archive.nytimes.com/bits.blogs.nytimes.com/2014/04/08/flaw-found-in-key-method-for-protecting-data-on-the-internet.

the reliability of forensic evidence in breath alcohol detection devices[147] and DNA analysis software.[148] Coding errors in 3D scanning and VCM software, for example, might result in distorted images that are relied upon by human firearm examiners to develop investigative leads or draw forensic conclusions used as evidence to wrongfully arrest, convict, and punish.

Product design errors, distinct from coding errors, present an additional source of error in automated forensic evidence and results, and are also documented across industries.[149]

Machines can malfunction, impacting forensic results. Machines naturally degrade over time. An old kitchen oven might not reach full temperature anymore; an old bathroom scale might read two pounds under accurate measure. Measurement devices require calibration to ensure continued consistency and accuracy. Routine and frequent testing can uncover degradation and allow for correction, but when left unattended, machines can malfunction in ways that manifest in skewed outputs.

Automation also introduces the risk of human machine operator error. Even automated tools rely on human operators to input data or evidence, implement settings, and determine thresholds. Machine operator error can alter results. An operator's incorrect placement of a bullet for 3D scanning might impact the resulting image, leading to the machine erroneously identifying non-existing similarities or dissimilarities between evidence items.

---

147. *In re* Source Code, 816 N.W.2d 525, 528, 543 (Minn. 2012) (affirming finding, after defense access, that a version of the source code has revealed errors impacting reliability of the Intoxilyzer, an instrument used to measure breath alcohol levels); State v. Chun, 943 A.2d 114, 120 (N.J. 2008) (finding, after defense access, that the source code for Alcotest 7110, an instrument used to measure breath alcohol levels, required "certain modifications . . . to permit its result to be admissible or to allow it to be utilized to prove a per se violation").

148. *See* David Murray, *Queensland Authorities Confirm "Miscode" Affects DNA Evidence in Criminal Cases*, COURIER MAIL (Mar. 20, 2015, at 22:00 PT), https://www.couriermail.com.au/news/queensland/queensland-authorities-confirm-miscode-affects-dna-evidence-in-criminal-cases/news-story/833c580d3f1c59039efd1a2ef55af92b (reporting that source code errors effecting a genotyping program called STRmix materially altered match statistics in DNA analysis in over sixty cases).

149. *See*, *e.g.*, OFF. OF DEFECTS INVESTIGATION ENF'T, NAT'L HIGHWAY TRAFFIC SAFETY ADMIN., INVESTIGATION REPORT 3, 16 (1978) (noting that the Ford Motor Company produced the Ford Pinto from 1970 to 1980, but its design had a serious design flaw that made the car prone to fuel leaks and explosions in rear-end collisions, resulting in 26 deaths and numerous injuries before a 1978 mass recall of the vehicle); Maribel Lopez, *Samsung Explains Note 7 Battery Explosions, and Turns Crisis into Opportunity*, FORBES (Jan. 22, 2017, at 22:31 ET), https://www.forbes.com/sites/maribellopez/2017/01/22/samsung-reveals-cause-of-note-7-issue-turns-crisis-into-opportunity (explaining that Samsung released the Galaxy Note 7 smartphone with a design error causing it to overheat and explode); *IKEA Reannounces Recall of MALM and Other Models of Chests and Dressers due to Serious Tip-Over Hazard; 8th Child Fatality Reported; Consumers Urged to Choose Between Refund or Repair*, U.S. CONSUMER PROD. SAFETY COMM'N, https://www.cpsc.gov/Recalls/2018/IKEA-Reannounces-Recall-of-MALM-and-Other-Models-of-Chests-and-Dressers-Due-to-Serious-Tip-over-Hazard (last visited Jan. 7, 2026) (reporting that one of Ikea's dressers was manufactured with a design flaw making it susceptible to tipping over easily, causing child deaths and injuries); *Peloton Recalls Tread+ Treadmills After One Child Died and More Than 70 Incidents Reported*, U.S. CONSUMER PROD. SAFETY COMM'N, https://www.cpsc.gov/Recalls/2021/Peloton-Recalls-Tread-Plus-Treadmills-After-One-Child-Died-and-More-than-70-Incidents-Reported (last visited Jan. 7, 2026) (reporting that Peloton created a treadmill with a design defect causing it to pull multiple children and pets under the machine, killing and injuring many).

Coding, product design, or machine errors can lead to incorrect conclusions used to wrongfully arrest, charge, convict, and punish. Lack of transparency can render these errors nearly impossible to expose.

E. LACK OF TRANSPARENCY SHIELDS AUTOMATION FROM SCRUTINY

Transparency in forensic methods is essential to enable scientific review,[150] protect rights to notice,[151] demand accountability,[152] and provide the public with some confidence in criminal legal system outcomes.[153] Automated technology is notoriously non-transparent. Private companies developing forensic tools for commercial gain have little incentive to reveal the workings of proprietary systems. These systems are sold to law enforcement agencies and firearms examiners. The use of automated tools for investigative purposes (rather than as direct evidence at trial) is often not disclosed at all.[154]

The need for transparency is heightened by the vulnerability of forensic tools to impure incentives. While some forensic disciplines have limited uses beyond law enforcement,[155] many do not. There is no commercial use for firearms examination tools beyond investigation and prosecution of gun-involved crimes, meaning creators of firearms examination technology have only one market: law enforcement.[156] The market success of a firearms examination tool can only be measured by its desirability to a law enforcement audience, which may quantify its success by troubling metrics such as arrests, charges, or convictions. This incentivizes tech developers to enable those results. Simply the specter of such an influence highlights the need for external, independent, and critical review of forensic technology.

There is no legal requirement for transparency in automated forensic tools.[157] Unlike other industries implementing automation, there is no regulatory

150. *See, e.g.*, Darrel C. Ince, Leslie Hatton & John Graham-Cumming, *The Case for Open Computer Programs*, 482 NATURE 485, 485 (2012) ("[A]nything less than release of actual source code is an indefensible approach for any scientific results that depend on computation . . . .").

151. *See* Patrick Toomey & Brett Max Kaufman, *The Notice Paradox: Secret Surveillance, Criminal Defendants, & the Right to Notice*, 54 SANTA CLARA L. REV. 843, 859–65 (2015) (articulating a due process right to notice of the use surveillance tools in a criminal case).

152. *See* Maneka Sinha, *The Dangers of Automated Gunshot Detection*, 5 J.L. & INNOVATION 63, 112 (2023) (calling for legislation and audits for the use of ShotSpotter technology as a mechanism for police accountability).

153. *See* Rebecca Wexler, *Liberty, and Trade Secrets: Intellectual Property in the Criminal Justice System*, 70 STAN. L. REV. 1343, 1402 (2018) ("[G]reater transparency could provide assurance that the outcome was proper [for anyone who is affected by a criminal justice system outcome]. At a minimum, this group includes defendants, victims, and their families and communities.").

154. *See supra* Subpart.III.B.3.

155. For example, DNA technology has both medical and commercial markets; fingerprint and facial recognition technology has commercial use.

156. This includes both direct law enforcement entities and private firearms examiners hired by law enforcement agencies. This author is unaware of, and was unable to locate, any documented non-law enforcement related, commercial use for firearms examination tools.

157. The need for a forensic regulatory agency was a major conclusion of the 2009 NRC report. NRC REPORT, *supra* note 25, at 16–17.

oversight. Medical technology impacts public health and therefore is subject to regulation by the Food and Drug Administration ("FDA"). Financial technology impacts consumer financial security and therefore is subject to regulation by the Consumer Financial Protection Bureau ("CFPB"). Forensic technology impacts the safety and liberty of individuals in our society and the integrity of the criminal legal system, and yet the government has not seen fit to regulate it. The National Institute of Standards and Technology ("NIST"), a government agency within the United States Department of Commerce, develops standards for various sectors, including the use of technology in criminal case work, but NIST is explicitly non-regulatory, so these standards are entirely optional.[158] NIST has no power to require that forensic practitioners implement their standards, nor any power to control what or how technology is used by police and prosecutors.

In the absence of a market or regulatory check on the validity of forensic technology, it is incumbent on developers to be transparent and invite independent review ensuring the validity of products impacting human liberty. Instead, they have done the opposite. This opacity contributes to the failures of the criminal legal system to address the scientific and legal implications of forensic automation.

F. AUTOMATION MAKES EXAMINATIONS MORE EFFICIENT

Automation offers one inarguable advantage over traditional methods of firearms examination: efficiency. A machine can work faster than a human; it can conduct multiple functions at one time,[159] and it does not need lunch breaks. Efficiency is more than simply a convenience; efficiency enables more comparisons to be conducted, allows resources to be redirected to other aspects of an investigation, and prevents slowing the criminal court process while awaiting forensic testing.

But efficiency is not always desirable in the criminal legal system. While efficient case resolutions can serve diverse interests (including those of the criminally accused), efficiency interests can also result in shortcuts to justice and due process. Inefficiency in criminal investigative techniques, causing law enforcement to work harder and use more resources, may encourage more careful decisionmaking in their investigations and prosecutions, ultimately reducing over-policing and over-prosecution.

158. *See National Institute of Standards and Technology*, FED. REG., https://www.federalregister.gov/agencies/national-institute-of-standards-and-technology (last visited July 24, 2025).

159. Laura Knowles, Daniel Hockey & John Marshall, *The Validation of 3D Virtual Comparison Microscopy (VCM) in the Comparison of Expended Cartridge Cases*, 67 J. FORENSIC SCIS. 516, 516 (2021) (finding that VCM is "appropriate and valid" for cartridge case comparison and routine casework, citing efficiency as a key benefit).

III. FAILURES OF THE CRIMINAL LEGAL SYSTEM IN ADDRESSING AUTOMATED FORENSICS

That our criminal legal system fails to adequately address the issues raised by automating forensic examinations should come as no surprise. Legal rules governing forensic evidence were designed for *human* forensic experts and *human* methods of examination. Having failed to sufficiently gatekeep even forensic evidence contemplated in their inception, these rules are clearly inadequate to address the novel features of automation. This Part surveys the ways in which the criminal legal system is *ill-suited* to resolve issues raised by automated forensic evidence due to the system's nature, structure, and actors, and how the system is *ill-equipped* to do so in individual cases due to structural, procedural, and practical barriers to meaningful access and adjudication.

A. CRIMINAL COURTS AS UNSUITABLE ARBITERS OF SCIENTIFIC VALIDITY

Rather than undergoing the rigors of actual scientific testing, forensic evidence has found validation in criminal courtrooms, which serve as both creators[160] and interpreters of questions around forensic validity. But criminal courts are unsuitable arbiters of scientific reliability. Following the imperative of *Daubert v. Merrell Dow Pharmaceuticals*[161] to consider scientific validity before admitting forensic evidence at trial, federal and some state courts began to take a closer look at forensic methods. Yet they continued to overwhelmingly admit forensic expert witnesses in criminal cases.[162]

Continued admission of unscientific evidence may flow in part from fundamental differences in law and science. While the interests of law and science may at times overlap, the interests in a legal dispute might also diverge from (or directly conflict with) unencumbered pursuits of scientific validation. In science, the question of scientific validity is the end in itself; in law, the same question is incidental to legal concerns, conditioned on the facts and procedural posture of a particular case,[163] and necessarily constrained by ethical norms as well as "subjective and normative concerns."[164]

160. *See* Sheila Jasanoff, *Law's Knowledge: Science for Justice in Legal Settings*, 95 AM. J. PUB. HEALTH S49, S53 (2005) ("[P]ost-*Daubert* judges have emerged as active participants in *making* science, consistent with their lay understandings of how science should be made.").

161. 509 U.S. 579, 593 (1993). *Daubert* changed the legal standard for admission of scientific testimony in federal courts, directing courts to directly address the scientific validity of expert testimony before allowed it at trial. Under the prior standard of *Frye v. United States*, still utilized in several states, courts need not consider scientific validity, but only whether the evidence was created by methods generally accepted in the relevant scientific community. 293 F. 1013, 1014 (D.C. Cir. 1923). The "relevant scientific community" is often defined to include only forensic practitioners.

162. Jim Hilbert, *The Disappointing History of Science in the Courtroom:* Frye*,* Daubert*, and the Ongoing Crisis of "Junk Science" in Criminal Trials*, 71 OKLA. L. REV. 759, 799–800 (2019).

163. *See* Jasanoff, *supra* note 160, at S51 ("Knowledge relevant to a legal proceeding is generated for the purpose of rendering justice within that specific setting. Scientific knowledge is also situated, but it is positioned chiefly so in relation to communities of theory and practice.").

164. *Id.* at S52.

Criminal courts are also inept scientific testers for more concrete reasons: The adversarial system discourages independent review, judges are not scientists, and the case and controversy requirement leaves the law lagging far behind technology.

The adversarial design of criminal courts disincentivizes, and in many cases precludes, independent scientific inquiry.[165] When considering a new method or tool in a criminal case, each party's primary interest is to advance their client's position; discerning the validity of novel technology may be incidental, but not a goal in itself. If novel DNA software renders incriminating results against the accused, the prosecution will be inclined to champion the technology, while the defense attorney will be interested in undermining it. Experts hired by each party are incentivized to support their (paying) client's position. The adversarial system discourages both sides from conducting open, unbiased scientific inquiries into validity and reliability. It also limits the judge's review of forensic reliability to two competing narratives. While "[s]cientific inquiry, too, may involve a choice between competing hypotheses, . . . these are not generally associated with questions of liability, blame, economic interest, or social justice,"[166] which undoubtedly impact decisions, whether intentionally or implicitly.

Additionally, the determiners of forensic admissibility—a decision requiring an examination under *Daubert* of scientific validity, if inadequate in practice[167]—are judges. Judges are highly educated, and some may be bright and thoughtful, but they are neither scientists nor subject matter experts in forensics. *Daubert* has been interpreted as requiring judges to "think like scientists,"[168] but demanding that they do so without scientific training or education. Faced with complex forensic decisionmaking, the path of least resistance is to simply admit evidence and lean on defense attorneys and jurors to uncover and discern scientific shortcomings. Judges, whether elected or politically appointed, might also face considerable political pressure to be tough on crime or rule favorably for the prosecution.[169] Not only does deferential admissibility screening defy the imperative to courts to act as "gatekeepers,"[170] it also simply does not work. Forensic evidence holds great power over jury decisionmaking, and even

165. *See* NRC REPORT, *supra* note 25, at 110 ("[T]he adversarial process relating to the admission and exclusion of scientific evidence is not suited to the task of finding 'scientific truth.'").

166. Jasanoff, *supra* note 160, at S52.

167. *Id.* at S53 ("In practice, however, *Daubert* and its progeny considerably widened the federal courts' maneuvering room with respect to admissibility, offering lower-court judges a broad and largely uncontrolled grant of discretion to declare case-by-case what counts as "science.").

168. *Id.* at S57 n.8 ("More accurately perhaps, the authors of the *Daubert* majority opinion asked trial court judges to think like scientists as *they* imagined scientists think.").

169. *See* Sanford C. Gordon & Gregory A. Huber, *The Effect of Electoral Competitiveness on Incumbent Behavior*, 2 Q.J. POL. SCI. 107, 108, 128, 133 (2007) (finding that judges facing noncompetitive retention elections sentence less severely than those facing partisan elections).

170. Jasanoff, *supra* note 160, at S51.

judicial attempts to limit the scope of forensic conclusions do little to temper the impact of junk science on jurors.[171]

The "case and controversy" clause[172] renders the law necessarily reactive, a poor fit for the dynamic nature of science and technology. Judges do not update legal interpretations to reflect advances in forensic evidence or understanding unless called upon to do so by a particular case, typically where an injustice has resulted from the prior interpretation. Countless comparable harms might occur in cases not reaching higher court review due to a plea agreement, dismissal, failure of an attorney to make an objection, or any number of reasons. The inherently reactive nature of law is simply unable to keep up with the quickly evolving nature of forensics and technology. The human costs are substantial; even a forward-looking judicial opinion will almost always come too late to provide meaningful, systemic protection. These structural features of the criminal legal system are particularly ill-suited to address the nature and pace of technological innovation.

B. BARRIERS TO MEANINGFUL REVIEW OF AUTOMATED EVIDENCE IN A CRIMINAL CASE

Criminal courts are ill-suited screeners of forensic validity, but until broader reforms take hold, they remain in this role. This places the primary task of exposing shortcomings of automated forensic machines on defense attorneys through the adversarial process. The Supreme Court has long recognized that the right to due process requires arming a defense team with the information and tools needed to mount an effective defense:

> [M]ere access to the courthouse doors does not by itself assure a proper functioning of the adversary process, and . . . a criminal trial is fundamentally unfair if the State proceeds against an indigent defendant without making certain that he has access to the raw materials integral to the building of an effective defense.[173]

When algorithms create evidence used or relied upon in a criminal prosecution, the underlying source codes controlling algorithmic function are "raw materials" integral to the building of an effective defense. Yet procedural and constitutional rules of discovery fail to directly contemplate the implications of automation on discovery and access rules, while additional barriers—obscurity of automation in the "shadows" of the investigation, weaponization of trade secrets protections, and difficulties in securing defense expert assistance—compound the access issues faced by defenders.

---

171. Nicholas Scurich, David Faigman & Brandon L. Garrett, *Ineffectiveness of the "Consistent With" Judicial Limitation on Forensic Firearm Identification Testimony*, 49 LAW & HUM. BEHAV. 387 ("[Limiting expert identification statements to] 'consistent with' did not significantly reduce guilty verdicts compared with definitive identification testimony, suggesting that it may not effectively convey limitations to jurors.").

172. U.S. CONST. art. III, § 2, cl. 1.

173. Ake v. Oklahoma, 470 U.S. 68, 77 (1985).

1. Discovery Rules Do Not Address Automation

Manufacturers of VCM publicize the technology's functions[174] but decline to disclose computer codes controlling those functions and resulting outputs. The threshold requirement to enable a defense challenge of a forensic tool or method used by the prosecution is access. For automated tools, this necessarily requires access to computer codes controlling those methods. Yet discovery of automated evidence remains hotly contested and unaddressed by discovery rules.

Discovery rules do not directly or clearly address the nature of machine-aided or created testimony, leaving ample room for disagreement among parties and confusion among courts regarding the application of discovery rules to algorithm source codes. Discovery rules vary significantly from jurisdiction to jurisdiction. Some states have adopted essentially "open file" discovery rules, in which prosecutors disclose all, or nearly all, evidence not protected by privilege to their adversary.[175] Other jurisdictions, including federal courts, severely restrict defense access to discovery until shortly before trial.[176] Most states take an intermediate approach.[177] While the specifics vary from jurisdiction to jurisdiction, discovery rules typically include specific regulations for the disclosure of expert reports and evidence to the defense. For example, a relevant section of North Carolina's discovery rule requires that prosecutors offering expert testimony disclose to the defense a "report of the results of any examinations or tests conducted by the expert," "the expert's curriculum vitae, the expert's opinion, and the underlying basis for that opinion."[178] The rule, similar to rules across many jurisdictions, assumes the expert opinion was drawn by a human, not a machine. Absent from the rule is mention of algorithms, source codes, validation studies, calibration, and maintenance records. While a litigant can argue for the logical application of the rule to these items and information, their opponent will likely disagree. Discovery rules fail to resolve critical issues of access.

2. Confrontation Clause Jurisprudence Does Not Contemplate Automation

Discovery disputes also implicate the Sixth Amendment's Confrontation Clause. The Confrontation Clause guarantees to the criminally accused the right to be "confronted with the witnesses against him."[179] The Confrontation Clause

174. Manufacturers feature their products' claimed attributes on their websites. *See, e.g.*, CADRE FORENSICS, *supra* note 96; LEADSONLINE, *supra* note 96.

175. Kenneth Williams & Richard G. Singer, *Chapter 6: Evidence Disclosure (Discovery*), in EXAMPLES & EXPLANATIONS, CRIMINAL PROCEDURE II: FROM BAIL TO JAIL 103–05 (5th ed. 2022).

176. *Id.*

177. *Id.*

178. N.C. GEN. STAT. § 15A-903 (2010).

179. U.S. CONST. amend. VI.

has been interpreted to only apply to evidence that is "testimonial"[180] in nature and to include a procedural right to cross-examination of that evidence.[181] Courts have clarified that confrontation rights extend to experts[182] and sworn affidavits presented in lieu of live testimony.[183] They have not, however, uniformly clarified confrontation rights in the context of machine testimony.[184]

Ensuring the reliability of evidence is a primary, if not ultimate, goal of the Confrontation Clause.[185] When a human witness provides incriminating testimony, their statements may be cross-examined and impeached to challenge or disprove their credibility. Machines and algorithms cannot be *cross-examined*, but they can be *examined* by experts. To satisfy the same constitutional imperative as cross-examination, an expert examination of an automated machine's "testimony" requires disclosure of the underlying source code. Source codes can include errors that impact function and output, resulting in unreliable evidence.[186] Such errors are impossible to detect without access to the source code.[187]

In addition to reliability interests, the Confrontation Clause implicates dignity interests. Automated evidence diminishes the moral accounting[188] of confrontation. As Professor Andrea Roth notes, "[p]erhaps it is easier to accuse someone when one builds an algorithm to do so."[189] The human implications of machine testimony are profound; algorithm-produced evidence is used as a tool to arrest, convict, punish, and separate families. Morally distancing these results from the actions of computer engineers, machine operators, and law enforcement

180. Although not clearly defined, "testimonial" statements seemingly include "out-of-court written or oral statements meant or understood to provide some form of evidence for use at trial, especially if made solemnly and to a state actor or agent." Ronald J. Coleman & Paul F. Rothstein, *A Game of* Katso *and Mouse: Current Theories for Getting Forensic Analysis Evidence Past the Confrontation Clause*, 57 AM. CRIM. L. REV. 27, 27 (2020).

181. Crawford v. Washington, 541 U.S. 36, 61 (2004).

182. Melendez-Diaz v. Massachusetts, 557 U.S. 305, 312 (2009); Bullcoming v. New Mexico, 564 U.S. 647, 665 (2011); Smith v. Arizona, 602 U.S. 779, 803 (2024).

183. *See Crawford*, 541 U.S. at 43–46 (ex parte affidavits implicate the core concerns underlying the Clause).

184. Confrontation Clause disputes over the disclosure of proprietary DNA analysis software have resulted in mixed and unsatisfactory results. "Although testability turns on disclosure of the source code, judges rarely order disclosure." Even when ordered, it "has not disclosed the code to the extent necessary for independent verification and validation." Natalie Murphy, *Give Me Liberty or Give Me the Source Code: Challenging a Black-Box Computer Algorithm Under* Daubert, 30 RICH. J.L. & TECH. 348, 396 (2024).

185. *Crawford*, 541 U.S. at 61.

186. *See supra* Part.II.

187. *See* A. Morin, J. Urban, P.D. Adams, I. Foster, A. Sali, D. Baker & P. Sliz*, Shining Light into Black Boxes*, 336 SCI. 159, 159 (2012) ("In the absence of source code, the inner workings of a program cannot be examined, adapted, or modified."); Christian Chessman, *A "Source" of Error: Computer Code, Criminal Defendants, and the Constitution*, 105 CAL. L. REV. 179, 183–99 (arguing that access to source code is necessary to prevent or unearth a number of structural programming errors); Erin E. Kenneally, *Gatekeeping Out of the Box: Open Source Software as a Mechanism to Assess Reliability for Digital Evidence*, 6 VA. J.L. & TECH., 2001, at ¶ 14 (same).

188. *See, e.g.*, Colin Allen, *The Future of Moral Machines*, N.Y. TIMES: OPINIONATOR (Dec. 25, 2011, at 17:30 ET), https://archive.nytimes.com/opinionator.blogs.nytimes.com/2011/12/25/the-future-of-moral-machines (noting issues with "battlefield machines").

189. Roth, *supra* note 132, at 2042.

agents itself erodes confrontation rights. Whether disclosure of source code alone is sufficient to satisfy the dignitary interests of the Confrontation Clause raises another open question. An added requirement of human confrontation may be necessary to satisfy the dignity interests of the Confrontation Clause,[190] but no human witness will substitute for disclosure of source code.

The Confrontation Clause demands access by the accused to their accuser. When that accuser is an algorithm or machine, the mechanism of confrontation may look different, but it is no less constitutionally compelled. Yet the application of Confrontation Clause rights to automated evidence remains unsettled, providing inadequate protection for the criminally accused.

3. Investigative Techniques Limit Scrutiny of Automated Forensics

Automated forensic tools have long evaded legal review. Rather than offer algorithm outputs as direct evidence at trial, prosecutors frequently limit—or obscure with "parallel reconstruction"[191]—their use to the shadows of the investigative stage, often allowing it to go undetected and unreviewed.[192]

Investigative techniques have allowed automated firearms examination tools to evade legal review. Earlier versions of the technology have been used by the FBI and ATF since the 1990s to image and compare evidence in a nationwide database, the National Integrated Ballistic Information Network ("NIBIN").[193] Police departments across the country can submit ballistic evidence to NIBIN for comparison with their database.[194] NIBIN relies on algorithms to create virtual images of submitted evidence, then to search the database for other evidence with similar toolmarks.[195] If there is a NIBIN "hit," meaning the newly entered projectile is determined by algorithms to have similar toolmarks to existing evidence in the database, a human firearm examiner will then view the images to confirm the machine's conclusion so that the conclusion can be used for court purposes.[196] Proponents claim this final step offers a human

190. *See generally* Ronald J. Coleman, *Human Confrontation*, 61 WAKE FOREST L. REV. (forthcoming 2026) (manuscript at 18–25), https://ssrn.com/abstract=5331746 (discussing how the Confrontation Clause should serve the value of dignity).

191. A practice where law enforcement conceals the true source of evidence by creating an alternative investigative pathway that appears to have led to the same discovery independently. *See generally* HUM. RTS. WATCH, DARK SIDE: SECRET ORIGINS OF EVIDENCE IN US CRIMINAL CASES (John Raphling, Cynthia Wong, Alison Parker, Dinah PoKempner & Joe Saunders eds., 2018), https://www.hrw.org/report/2018/01/09/dark-side/secret-origins-evidence-us-criminal-cases (describing how parallel reconstruction renders invisible evidence obtained through illegal operations).

192. *See* Lau, *supra* note 18 (manuscript at 41).

193. *See* BUREAU OF ALCOHOL, TOBACCO, FIREARMS & EXPLOSIVES, *supra* note 18.

194. *Id.*

195. *See* ROBERT C. TROYER, NIBIN TOOLKIT FOR PROSECUTORS 17 (2d ed. 2024), https://46525538.fs1.hubspotusercontent-na1.net/hubfs/46525538/Product%20documentation/LO_NIBIN-ToolkitForProsecutors_20240722_WEB.pdf; William King, Charles Katz, Edward Maguire & James Frank, *Opening the Black Box of NIBIN: A Descriptive Process and Outcome Evaluation of the Use of NIBIN and Its Effects on Criminal Investigations, Executive Summary*, U.S. DEP'T OF JUST. OFF. OF JUSTICE PROGRAMS 1 (2013), https://www.ojp.gov/pdffiles1/nij/grants/243977.pdf.

196. TROYER, *supra* note 195, at 13; King et al., *supra* note 195, at 1.

check on the technology; critics identify it as a tool to shield the real decision-maker—the machine—from disclosure and scrutiny. Adding human confirmation allows prosecutors to provide discovery related to the human confirmation process and conclusion, rather than the automated NIBIN hit that actually developed the lead. This practice, known as "parallel reconstruction,"[197] is used across different forms of investigation to shield investigative activity from review, whether to mask the discovery of evidence or to protect a novel forensic tool from reliability review.

That VCM and automated forensic technology has evaded review for so long is especially remarkable given the central role it plays in criminal prosecutions. NIBIN's website boasts that the database stores seven million pieces of ballistic evidence and has generated 1,150,000 leads during its twenty-seven-year history, including 217,000 leads in 2024 alone.[198] Many local police departments have purchased their own automated forensic tools; how they are using them is largely unknown. Public defenders in jurisdictions where police departments are known to possess 3D scanners or VCM machines have expressed their belief that the technology is not being used in case work.[199] More likely, their use is obscured and undisclosed. Investigative techniques like parallel reconstruction have successfully erected barriers to access and review of automated tools.

4. Trade Secrets Rights Limit Scrutiny of Automated Forensics

Private forensic companies refuse to disclose proprietary source codes,[200] asserting legal trade secrets rights. In many instances, they further argue that access to the source codes and algorithms themselves is not necessary for defense attorneys to view, but rather they need only understand the machine's "basic principles" to mount an adequate defense.[201] Researchers disagree, explaining, "[c]ommon implementation errors in programs . . . can be difficult to detect without access to source code."[202]

The United States Patent and Trademark Office defines a trade secret as: "information that has either actual or potential independent economic value by virtue of not being generally known," that "derives value from disclosure or use of the information by others who cannot ascertain the information through proper means" and "is subject to reasonable efforts to maintain [its secrecy]."[203]

197. *See supra* note 191.

198. *See National Integrated Ballistic Information Network*, BUREAU OF ALCOHOL, TOBACCO, FIREARMS & EXPLOSIVES (2025), https://www.atf.gov/resource-center/fact-sheet/2024-national-integrated-ballistic-information-network.

199. This statement is based on direct inquiries to public defenders' officers.

200. I am unaware of an instance in which a prosecutor has joined defense request for source codes, though I imagine an ethical prosecutor concerned with questions of forensic integrity might be so inclined.

201. Roth, *supra* note 132, at 2028.

202. *Id.* at 2028 (citation omitted); Morin et al., *supra* note 187.

203. *Trade Secret Policy*, U.S. PATENT & TRADEMARK OFF., https://www.uspto.gov/ip-policy/trade-secret-policy (last visited July 25, 2025) ("All three of the listed elements are required. If any one of them ceases to

If something is designated as a "trade secret," courts can, among other protections, order that it be shielded from public disclosure.[204] Because the trade secret designation depends on it not being generally known, a single disclosure[205] of a company's source codes would compromise future assertion of trade secrets rights by the company. Accordingly, a forensic company has an interest in vehemently opposing disclosure of its algorithm source codes in every case simply for its own economic benefit.

The application of trade secret rights to issues of discovery in criminal cases has been considered by academics[206] and courts alike. Trade secrets rights appear in direct conflict with the accused's constitutional rights to compel access to source codes.

Many courts have wrestled with the tension between trade secret rights and confrontation/discovery rights in the context of automated DNA probabilistic genotyping software. Like VCM, DNA probabilistic genotyping software replaces human forensic examiners with algorithmic machine processes.[207] DNA probabilistic genotyping software is used to conduct comparison testing on evidence containing a mixture of the DNA of multiple contributors,[208] which is more complex than one-to-one DNA comparisons.[209] Unlike VCM, the conclusions of DNA probabilistic genotyping software are routinely offered in court as direct evidence, and the software has been subject to numerous rigorous legal challenges. Defense requests for pretrial disclosure of the source codes

---

exist, then the trade secret will also cease to exist. Otherwise, there is no limit on the amount of time a trade secret is protected.").

204. U.S. PATENT & TRADEMARK OFF., OFF. POL'Y & INT'L AFFS., THE DEFEND TRADE SECRETS ACT AT FIVE: THE INEVITABLE DISCLOSURE DOCTRINE (2021), https://www.uspto.gov/sites/default/files/documents/USPTO-DefendTradeSecretsAct-atFive.pdf.

205. "[A] trade secret can lose its protected status if it is disclosed . . . either through legal filings . . . or through accidental or intentional disclosure by an employee . . . . At least one court has held that information can lose its status as a trade secret through an anonymous posting on the Internet, even for a very limited time." U.S. Dep't of Just., Crim. Res. Manual § 1127 (n.d.) (internal citations omitted). This can be problematic for companies because criminal proceedings are generally publicly accessible. *See, e.g.*, *Access to Court Proceedings*, U.S. COURTS, https://www.uscourts.gov/court-records/access-court-proceedings (last visited July 25, 2025). Nevertheless, private information in case files can be protected through redacted filings and court orders sealing certain documents containing confidential or classified information. *See id.*

206. *See* Rebecca Wexler, *Life, Liberty, and Trade Secrets: Intellectual Property in the Criminal Justice System*, 70 STAN. L. REV. 1343, 1403–07 (2018).

207. DNA probabilistic genotyping software ("PGS") automates genetic analysis by using computer simulations to compare DNA profiles and calculate probabilistic matches between evidence samples and persons of interest, replacing traditional human examination in complex scenarios involving degraded samples or multiple DNA contributors. Approximately 100 U.S. laboratories use PGS, including the FBI, which began using the STRmix software package in 2015. U.S. GOV'T ACCOUNTABILITY OFF., GAO-19-707SP, SCIENCE AND TECH SPOTLIGHT: PROBABILISTIC GENOTYPING SOFTWARE (2019).

208. *See* Decision & Order on DNA Analysis Admissibility at 7 n.7, People v. Hillary, Indictment No. 2015-15 (N.Y. Cnty. Ct. St. Lawrence Cnty. Ct. Aug. 26, 2016).

209. *See When DNA Samples Are Complicated: Calculating Variation in Mixed Samples Interpretation*, NAT'L INST. OF JUST. (Dec. 12, 2022), https://nij.ojp.gov/topics/articles/when-dna-samples-are-complicated-calculating-variation-mixed-samples-interpretation.

controlling these DNA genotyping algorithms have been met with mixed success in court.[210]

Many courts have flatly denied defense access to DNA software source codes.[211] Others have granted access, but with substantial limitations. For example, in *State v. Pickett*,[212] New Jersey's Appellate Division weighed the competing interests and ordered disclosure of the DNA software's source code to the defense, under a protective order, for review by their own expert.[213] The court wrote:

> As technology proliferates, so does its use in criminal prosecutions. Courts must endeavor to understand new technology—here, probabilistic genotyping—and allow the defense a meaningful opportunity to examine it. Without scrutinizing its software's source code—a human-made set of instructions that may contain bugs, glitches, and defects—in the context of an adversarial system, no finding that it properly implements the underlying science could realistically be made. Consequently, affording meaningful examination of the source code, which compels the critical independent analysis necessary for a judge to make a threshold determination as to reliability at a *Frye* hearing, is imperative.[214]

This decision recognized the need for gatekeeping as forensic technology evolves and the imperative of disclosure to protect the accused's constitutional rights.[215] *Pickett* did not, however, recognize an automatic or complete right to discovery of proprietary source codes; instead, the court premised disclosure on a showing of "particularized need"[216] by the defense and severely limited the

---

210. *Compare* State v. Pickett, 246 A.3d 279, 279 (N.J. Super. Ct. App. Div. 2021) (granting defendant access to TrueAllele source code under protective order after finding "particularized need" demonstrated), *and* People v. Superior Ct., No. B258569, 2015 Cal. App. LEXIS 105, at *8 (Jan. 9, 2015) (depublished) (ordering disclosure of TrueAllele source code under protective order to avoid "work[ing] injustice" on defendant's confrontation rights), *with* People v. Wakefield, 38 N.Y.3d 367, 385–86 (2022) (denying access to TrueAllele source code where defendant failed to demonstrate particularized need and holding that source code disclosure was not required under the Confrontation Clause because source code is not testimonial and cannot be cross-examined as a declarant), *and* Decision & Order on DNA Analysis Admissibility, *supra* note 208, at 7–8 (finding STRmix generally accepted but inadmissible due to lab's lack of internal validation, without addressing source code disclosure). *See also* Commonwealth v. Skundrich, 327 A.3d 218, 222 (Pa. Super. Ct. 2024) (finding potential due process violation in failing to provide TrueAllele source code access and remanding for evidentiary hearing).

211. *See, e.g.*, *Wakefield*, 38 N.Y.3d at 386 (denying access to TrueAllele source code); State v. Loomis, 881 N.W.2d 749, 753 (Wis. 2016); (denying defendant access to COMPAS algorithm source code used in sentencing risk assessment); State v. Ghigliotty, 232 A.3d 468, 486 (N.J. Super. Ct. App. Div. 2020) (denying access to BULLETTRAX algorithm source code).

212. *Pickett*, 246 A.3d at 278.

213. *Id.*

214. *Id.* at 323–24.

215. *Id.* at 278.

216. "In summary, defendant articulated a particularized need for the proprietary source code and related information for use at the *Frye* hearing by (1) demonstrating a rational basis for ordering the State to attempt to produce it, including through expert testimony supporting the claim for disclosure; (2) providing specificity for the information sought; (3) showing through examples from other jurisdictions that the company's intellectual property can be safeguarded by a protective order; and (4) demonstrating that source-code review is particularly crucial to evaluating the unique technology at issue here." *Id.* at 324.

scope of the defense expert's access by protective order. These legal requirements, tailored to balance competing legal rights of the accused and private companies, in practice, present unwieldy barriers to meaningful review. Overbroad protective orders are an inadequate solution to the tensions of commercial and constitutional rights.

Demonstrating a "particularized need" to overcome trade secrets interests remains an insurmountable barrier absent a defense expert with subject matter expertise. This creates a Catch-22 for defense attorneys: Without access to how an automated tool works, outside experts cannot develop expertise, but without a defense expert, courts will not compel discovery demonstrating how it works.

5. Unavailability of Defense Experts Limits Scrutiny of Automation

Just as "mere access to the courthouse doors does not by itself assure a proper functioning of the adversary process,"[217] neither does mere access to an algorithm's source code. Even if forensic companies were inclined or court-ordered to hand over source codes, absent a relevant computer science background, defense attorneys would be unlikely to glean their import. Given the technical nature of machine evidence, the "raw materials" integral to the building of an effective defense include both the source codes *and* a qualified expert to review and interpret them.

There are many practical barriers to accessing a qualified defense expert for novel automated technology. Defenders, especially public and court-appointed defenders, face resource limitations. But even with adequate funding, defenders face challenges in retaining experts to confront automated forensic evidence. The limited product market[218] limits opportunities for independent computer programmers or product developers to gain particularized subject matter expertise. For example, in *Ghigliotty*, a nationwide search[219] revealed two categories of experts then familiar with VCM. The first included individuals involved in developing similar technology themselves, and therefore directly interested in preserving legal protections of proprietary code and technology, contrary to the interest of Mr. Ghigliotty. The second was a smaller group of individuals working to create standards to regulate the development of VCM; they were, at the time, unwilling to work with defense teams challenging the technology because they wanted to remain unbiased. Absent an available expert with expertise in both computer science *and* the algorithmic tool at issue, litigants face an uphill battle.[220] Structural forces—limited market, legal

217. Ake v. Oklahoma, 470 U.S. 68, 77 (1985).

218. *See supra* Subpart.II.F.

219. This search included inquiries with expert databases, firearms examination experts, defense attorneys and other litigants experienced in challenging firearms examination evidence and forensic evidence, other developers of automated firearms examination technology, and statisticians and researchers.

220. Uphill does not mean impossible. Defense attorneys can hire experts in computer programming, automation, or forensic issues more broadly to mount challenges—such as a computer science expert to address

protections for proprietary software, and the disinclination of forensic technology companies to respond to calls for transparency—inhibit development of such an expert.

Importantly, a defense expert may help individual litigants, but they cannot substitute for independent scientific validation testing, and will not alone guarantee meaningful review. Protective orders limiting the scope, time, or space of source code access[221] undermine, or render impossible, meaningful scientific review. Researchers calling for open access write that independent testing and validation requires "automated testing over thousands of samples,"[222] which is impossible under these constraints.

Discovery rules, investigative techniques, trade secret rights, and defense expert unavailability often erect insurmountable barriers to defense attorneys even getting their hands on the raw materials needed to challenge automated forensic evidence. If they do, they face the next set of hurdles, navigating admissibility and credibility challenges.

C. BARRIERS TO ADJUDICATING AUTOMATED FORENSIC EVIDENCE

Evidentiary rules governing admissibility and use of evidence in criminal trials were written before automation entered the landscape of criminal prosecutions and have yet to be adequately modified to provide necessary guidance for judges or tools for litigants in adjudicating admissibility and credibility questions. When courts are called upon to decide the applicability of existing rules to novel technology, the questions raised often go unanswered due to the nature of a criminal legal system that resolves the vast majority of cases through plea bargains, preventing the creation of guidance or protections for the next case.

1. Admissibility Rules Fail to Address Automation

While the threshold[223] for admission of evidence is very low—typically, anything "relevant"[224] and "authentic"[225] is fair game—additional requirements are placed on forensic evidence. But questions of *which* and *how* these existing evidence rules apply to automated forensic evidence remain open, failing to offer litigants and courts any meaningful guidance.

---

programming errors and a statistician to address foundational validity, etc. Doing so will require more resources and more research and work by the attorney to tie the pieces together for the judge.

221. *See supra* Subpart.III.B.4.

222. Michael D. Edge & Jeanna Neefe Matthews, *Open Practices in Our Science and Our Courtrooms*, 38 TRENDS GENETICS 113, 113–14 (2022).

223. A number of other evidentiary rules may come into play depending on the specific evidence, facts, and legal issues. *See infra* notes 224–228.

224. Evidence is "relevant" if it has "any tendency to make a fact [of consequence in determining the action] more or less probable than it would be without the evidence." FED. R. EVID. 401.

225. "Authenticity" simply requires a showing that the evidence is "what it purports to be." FED. R. EVID. 901.

Forensic evidence is subject to admissibility rules governing expert testimony: *Daubert*,[226] *Frye*,[227] Federal Rule of Evidence 702,[228] or a local equivalent. While *Daubert*, *Frye*, and Rule 702 hearings apply different standards (*Daubert* and Rule 702 require consideration of scientific validity,[229] while the exceedingly more deferential *Frye* simply looks for acceptance by the relevant scientific community),[230] all theoretically provide *some* protection against the admission of demonstrably unreliable forensic methods but lack sufficient scrutiny.

Even the more exacting *Daubert* standard, which contemplates admission of forensic evidence through a human expert, is not explicitly applicable to machine methods[231] and cannot be readily superimposed onto them. Adding further complication, analysis of a human expert's conclusion *aided* by automation, as in *Ghigliotty,* will necessarily differ from that of machine output offered as evidence; both the human and machine roles must be scrutinized. But when a human expert uses automated tools in their examination, *Daubert* scrutiny has focused on methods used by the human examiner, as opposed to the machine, even in instances where the human examiner acts as the "mere scrivener"[232] for the machine's output or recreates the machine's output through "parallel reconstruction."[233] This approach treats VCM no different than its predecessor, the confocal microscope. But while a microscope merely magnifies evidence, VCM creates and manipulates evidence. This distinction is critical and illustrative of the inadequacy of the existing *Daubert* framework to address automation-aided expert testimony.

Even where machine output is itself offered as the expert conclusion, as with DNA probabilistic genotyping software, evidence is typically entered through a human witness. A machine operator testifies regarding inputs (what evidence or images were submitted for comparison), settings or functions chosen by the operator, and authenticity—distinct from accuracy[234]—of the machine output (comparison conclusion). But machine operators cannot speak to the methodology of the automated analysis.

---

226. *See* Daubert v. Merrell Dow Pharms., Inc., 509 U.S. 579, 579–80 (1993) (requiring judges to determine that scientific or technical methods underlying expert testimony be scientifically valid).

227. *See* Frye v. United States, 293 F. 1013, 1014 (D.C. Cir. 1923) (holding that novel scientific methods must "have gained general acceptance in the particular field in which [they] belong[]").

228. Rule 702, amended following *Daubert*, governs admissibility decisions in federal courts and requires, among other criteria, that the expert testimony: "is based on sufficient facts or data"; "is the product of reliable principles and methods"; and "reflects a reliable application of the principles and methods to the facts of the case." FED. R. EVID. 702.

229. *Daubert*, 509 U.S. at 579–80.

230. *Frye*, 293 F. at 1014.

231. Roth, *supra* note 132, at 2032.

232. "Under current law, courts treat the machine as the method of a human expert, rather than as the expert itself, even when the expert is a 'mere scrivener' for the machine's output." *Id.*

233. For a discussion of "parallel reconstruction," see *infra* Subpart.II.C.3.

234. *See* Edge & Matthews, *supra* note 222.

While expert evidence rules have technically not applied to machine conclusions absent a human witness, rule makers have recently worked to change that in federal courts. On June 10, 2025, the Committee on Rules of Practice and Procedure—the advisory body to the Judicial Conference of the United States—approved a new Federal Rule of Evidence, Rule 707, to address growing concerns around the use of AI-generated evidence. The rule states:

> Where machine-generated evidence is offered without an expert witness and would be subject to Rule 702 if testified to by a witness, the court must find that the evidence satisfies the requirements of Rule 702 (a)-(d). This rule does not apply to the output of basic scientific instruments.[235]

While a modest step in the right direction, Rule 707 still fails to adequately address the novel concerns raised by automation. Neither Rule 702 nor 707 directly accounts for, or demands courts to explicitly consider, the sources of potential error and bias arising from the nature of automation.[236] While Rule 707 establishes an admissibility test for automated forensic evidence, it fails to provide guidance as to what that review demands in practice. The rule also fails to define "simple scientific instruments," begging for litigation as to the scope of its applicability—again, without direction for the adjudicators, who themselves lack scientific or forensic expertise.

The majority of criminal cases are prosecuted in state courts; whether states will follow the federal courts' lead remains to be seen. Regardless, if reliability screening under Rule 707 looks anything like that under Rule 702 and *Daubert* (and certainly *Frye*), there is little cause for celebration. *Daubert,* Rule 702, and *Frye* have historically failed to ensure adequate gatekeeping for the forensic evidence they were designed to assess; application of the same rules to automation will surely bear the same or worse[237] results.

Even applying the more rigorous *Daubert* test, judges tend to admit evidence so long as the principles underlying the machine's methodology are sound and when validation studies appear to demonstrate low error rates.[238] But theoretical soundness does not guarantee accuracy in actuality, and error rates can be deceptive. Error rate studies in forensics are notoriously rife with problematic design—among other design flaws,[239] they are conducted under idealized conditions not representative of actual case work and are not conducted by independent, uninterested parties.[240]

---

235. Frank Young, *Propose New FRE 707*, UNIV. OF ILL. CHI. L. LIBR. (July 3, 2025), https://library.law.uic.edu/news-stories/proposed-new-fre-707.

236. *See supra* Subpart.II.E.

237. Outcomes might be worse, permitting *more* unreliable evidence, because Rules 702 and 707 do not direct courts to address the novel sources of error and bias posed by automation, forecasting an incomplete screening process. *See* FED. R. EVID. 702, 707.

238. Roth, *supra* note 132, at 1982, 2033 (discussing flaws in error rate studies, and explaining how error rate studies may provide inadequate bases to declare a machine outcome "accurate" due to issues including feedback loops and the absence of satisfactory metrics for testing forensic outcomes).

239. Cuellar, et al., *supra* note 87.

240. *See* United States v. Tibbs, 2019 D.C. Super. LEXIS 9, at *39 (D.C. Super. Ct. Sep. 5, 2019).

Overreliance on error rate studies for admissibility decisions is problematic. Proficiency tests cited to support validity routinely artificially lower error rates by throwing out "inconclusive" responses entirely when reporting results.[241] Aware they are being tested. . .

> Examiners (so the argument goes) . . . may default to saying inconclusive on difficult cases rather than reach an erroneous source conclusion, thereby mask[ing] what would be a mistaken identification or elimination in casework and substantially reduc[ing] the credibility and reliability of the error rates reported.[242]

Further, error rate studies are typically designed by firearm examiners lacking expertise in research science or study design.[243] Unsurprisingly, these studies "often fail to account for test-taking bias or to control for variation in test-takers' approaches to study problems,"[244] limiting their import to any meaningful scrutiny and providing a misleading assessment of scientific validity.

The application of these existing standards and practices to automation, as contemplated by Rule 707, forecasts unjust results. Cadre Forensic Technologies, a leading manufacturer of VCM technology, boasts that its tools "have been validated via peer reviewed studies and studies conducted by the FBI and Canadian RCMP," which demonstrate that Cadre-VCM is "at least as good and likely better than traditional light comparison microscopy," "showing a lower inconclusive rate."[245] But Cadre's VCM tool has only been tested in a limited way. There have been four published studies that aim to evaluate its method.[246] None were conducted by independent administrators. Three studies

241. Richard E. Gutierrez, *Bowling with Bumper Rails: How Firearms Examiners Have Duped the Courts and Generated Low Error Rates Only by Avoiding Challenging Comparisons*, 75 U.C.L.J. 1535, 1551–52 (2024).

242. *Id.* at 1551 (internal quotations omitted); Alan H. Dorfman & Richard Valliant, *Inconclusives, Errors, and Error Rates in Forensic Firearms Analysis: Three Statistical Perspectives*, FORENSIC SCI. INT'L: SYNERGY, 2022, at 7; *see* Itiel E. Dror & Nicholas Scurich, *(Mis)use of Scientific Measurements in Forensic Science*, 2 FORENSIC SCI. INT'L: SYNERGY 333, 337 (2020) (noting that "error rate studies fall short[] and produce inaccurate and misleading error rate estimates" when they do not account for correctness or incorrectness of inconclusive conclusions).

243. Sinha, *supra* note 25, at 930–31; *Tibbs*, 2019 D.C. Super. LEXIS, at *41; *see generally* Dror & Scurich, *supra* note 242 (explaining how testimony of firearm experts regarding error rates coupled with study design flaws can mislead jurors as to its accuracy).

244. Sinha, *supra* note 25, at 930.

245. *Validated Virtual Comparison Microscopy (VCM): Cadre's Validated VCM Tools Lead the Industry*, CADRE FORENSICS, https://www.cadreforensics.com/VirtualComparisonMicroscopy.html (last visited July 25, 2025).

246. Todd J. Weller, Marcus A. Brubaker, Pierre Duez & Ryan Lilien, *Introduction and Initial Evaluation of a Novel Three-Dimensional Imaging and Analysis System for Firearm Forensics*, 47(4) AFTE J. 198 (2015); Pierre Duez, Todd Weller, Marcus Brubaker, Richard E. Hockensmith II & Ryan Lilien, *Development and Validation of a Virtual Exam Tool for Firearms Forensics*, 63(4) J. FORENSIC SCI. 1069 (2018); Chad Chapnick, Todd J. Weller, Pierre Duez, Eric Meschke, John Marshall & Ryan Lilien, *Results of the 3D Virtual Comparison Microscopy Error Rate (VCMER) Study for Firearm Forensics*, 66 J. FORENSIC SCI. 557 (2020); Laura Knowles, Daniel Hockey & John Marshall, *The Validation of 3D Virtual Comparison Microscopy (VCM) in the Comparison of Expended Cartridge Cases*, 67 J. FORENSIC SCI. 516 (2021).

were conducted by representatives from Cadre Forensics,[247] and the fourth by the Canadian Royal Mounted Police, a law enforcement agency trying to validate a new method their lab acquired, and therefore holding a financial and professional stake in its validation.[248] The studies feature limited numbers of participants and samples, focus on matching evidence samples (meaning that the results are less indicative of the ability to exclude), and feature low or unknown levels of complexity.[249] No studies have permitted access to the proprietary algorithm source codes for outside review, and all lack the independence, process, and demonstration of repeatability necessary for meaningful scientific testing. These studies share methodological flaws with the error rate studies that have long been relied upon to support traditional firearms examination methods, and since exposed to be scientifically unsound.[250] Yet courts consistently defer to similarly flawed error rate studies in admissibility determinations.[251]

Compounded by the barriers to access and meaningful review of algorithms, defense attorneys are left to litigate admissibility without full information while facing antiquated rules ill-suited to the nature of algorithmic forensic technology. Admissibility challenges to automated firearms examination tools remain few and far between,[252] and they are often terminated by plea agreement before completion,[253] but lessons from courts' historical gatekeeping failures illustrate the need for a different approach.

2. Credibility Testing Does Not Account for Automation

When admitted at trial, forensic evidence is subject to testing through cross-examination and juror assessment. Legal rules, practices, and instructions guide litigants and jurors through these processes. Cross-examination and impeachment allow lawyers to test and expose weaknesses in the testimony of

247. *See id.* The "Weller," "Duez," and "Chapnick" studies all include as a co-author, either Todd Weller, a "Domain Expert" at Cadre Forensics, and/or Ryan Lilien, the "Founder and Head of Development" at Cadre Forensics. *Team*, CADRE FORENSICS, https://www.cadreforensics.com/About.html (last visited July 25, 2025).

248. *See* Knowles et al., *supra* note 246, at 517. The "Knowles" study was conducted by the Canadian Mounted Police, which had recently purchased Cadre's TopMatch-3D scanner and associated virtual comparison software and were seeking to validate it for their use in case work. Notably, VCM equipment is expensive. While the cost is not publicly available, conversations during the *Ghigliotty* litigation suggested acquiring BULLETTRAX would cost a police department several millions of dollars. Agencies making, or requesting, such funding have a financial and professional stake in its validation, undermining the independence of review.

249. *See supra* note 246 (collecting the four studies that evaluate the VCM method).

250. *See* Cuellar et al., supra note 87, at 2 (noting the following methodological flaws in firearms examination error rate study designs, which render their results invalid: inadequate sample size, non-representative sample, non-representative testing conditions and environment, nonconclusive responses are treated as correct or ignored, invalid or nonexistent uncertainty measures for error rates, and missing data).

251. Roth, *supra* note 132, at 1982.

252. *See supra* Subpart.III.B.

253. This author engaged in both legal research and extensive inquiries of defense attorneys and forensic specialists across the country to explore all instances in which defense attorneys have challenged the admissibility of algorithmic firearms examination methods. The only challenges discovered have been terminated before resolution of the legal challenge. In these cases, prosecutors have extended unusually lenient plea offers in lieu of seeing the litigation through.

adverse witnesses. Model jury instructions offer jurors guidance on how to assess the credibility of witnesses and evidence. These tools, again designed with human witnesses in mind, fail to provide meaningful tools for lawyers to challenge automated evidence and fail to provide jurors adequate guidance on how to assess it.

Defense attorneys cannot cross-examine a computer code or a machine. They *can* cross-examine human experts called to testify about the machine's function and methods, but the same barriers to meaningful exposure of machine methods present at *Daubert* hearings are replicated during trial. Lack of access to source codes, product design, and machine maintenance logs limit the ability to challenge algorithm-aided conclusions. Unavailability of,[254] or lack of access to,[255] experts to assist in preparing a cross-examination further undermines its efficacy. Absent a defense expert's explanation, most jurors lack the background to understand the import of a cross-examination designed to reveal shortcomings in an algorithmic method.

Defense attorneys can discredit adverse witness testimony through impeachment using the declarant's prior inconsistent statements, evidence of incapacity, bias, or character of dishonesty.[256] But what does impeachment of a machine look like? Others have suggested it might include scrutiny of the "character or capacity of human programmers, inputters, and operators."[257] An attorney may impeach the bias, credibility, or accuracy of a human expert involved in the development, operation, or interpretation of automated tools, but these are limited to the machine design and human use. Impeachment of the machine function itself requires access to source codes and product designs, as well as prior "statements" (machine outputs). Even with full access to this information, procedural rules and practices of the courtroom lack a clear process for exposing impeachable shortcomings of automated machines.

In theory, one might impeach a machine's reliability with known error rates. But error rate studies, notorious for poor design and misleading claims,[258] lack meaningful value for impeachment.[259] Further, even when presented with troubling error rates, research has shown a limited impact on jury decision-making. Jurors tend to interpret the importance of error rates through the lens of preexisting beliefs about a discipline's reliability.[260] If a juror is already suspicious of a forensic field, they will give more weight to error rate evidence; if a jury already believes in the field's reliability, they will be more likely to

254. *See supra* Subpart.III.B.5.

255. Experts are often financially inaccessible to the criminally accused represented by appointed or retained counsel.

256. *See, e.g.*, FED. R. EVID. 806.

257. Roth, *supra* note 132, at 2036.

258. *See supra* Subpart.III.C.1 (discussion of error rate studies).

259. *Id.*

260. *See* Brandon L. Garrett, William E. Crozier & Rebecca Grady, *Error Rates, Likelihood Ratios, and Jury Evaluation of Forensic Evidence*, 65 J. FORENSIC SCI., 2020, at 2.

disregard error rates in determining the credibility of the forensic testimony.[261] This illustrates how confirmation bias[262] shapes not only the creation of forensic evidence but also its assessment and application. These findings also reinforce the need for more fastidious judicial gatekeeping of forensic evidence.

Ultimately, jurors decide evidentiary credibility and value in criminal trials. They are given guidance on how to assess human credibility, including expert witnesses, by considering factors such as a witness's demeanor, interest in the outcome of the case, or biases, inconsistencies of the testimony with other evidence, and reasonableness of the testimony.[263] There are no comparable jury instructions that provide guidance for weighing the credibility of machines, which are also capable of making mistakes or providing outputs that are incorrect, misleading, or easily misinterpreted. The exoneration of innocent people convicted by juries relying on forensic evidence illustrates that juror scrutiny has failed to correctly distinguish between reliable and unreliable forensics.[264] We cannot expect a new generation of jurors to accurately distinguish between reliable and unreliable forensic technology, absent tailored tools for testing and judging credibility. Neither is presently provided.

3. The Coercive Plea System Prevents Resolution of Legal Issues

Legal questions arising from automation have been raised in criminal courts but remain unanswered. One reason for this is the nature of the plea system. Often when judges are called upon to make decisions about novel technology, as in *Ghigliotty,* prosecutors extend favorable plea offers that effectively render those decisions—which threaten to expose proprietary interests, demand more process, and uncover sources of machine unreliability—moot. Compounding these concerns, the coercive nature of draconian sentencing laws makes a guilty plea the rational choice for most of the criminally accused, regardless of guilt or innocence.[265] When a guilty plea cuts short litigation of novel technology, it can be incorrectly interpreted or presented as evidence of the new technology's legitimacy. Through this process, novel forensic tools

261. *Id.* at 5.

262. Confirmation bias exists where "individuals interpret information, or look for new evidence, in a way that conforms to their pre-existing beliefs or assumptions." *See* PCAST REPORT, *supra* note 34, at 31.

263. For example, a New York model jury charge instructs jurors that, when determining whether to credit a witness's testimony, they may consider: the witness's opportunity to observe the events to which they testified, recall ability, plausibility of the testimony, consistency or inconsistency with other evidence, the manner in which they testified, the witness's background and experience, bias, hostility, attitude, motive, interest/lack of interest, and previous criminal conduct. *Credibility of Witnesses*, N.Y. STATE UNIFIED CT. SYS., https://www.nycourts.gov/judges/cji/1-General/CJI2d.Credibility.pdf (last visited July 24, 2025).

264. *See* THE INNOCENCE PROJECT, *supra* note 32 ("The misapplication of forensic science is the second most common contributing factor to wrongful convictions.").

265. JED S. RAKOFF, WHY THE INNOCENT PEOPLE PLEAD GUILTY AND THE GUILTY GO FREE: AND OTHER PARADOXES OF OUR BROKEN LEGAL SYSTEM 20 (Farrac, Straus & Giroux eds., 1st ed. 2021).

evade legal review, while continuing to advance investigations, arrests, prosecutions, and coerced—functionally[266] if not legally[267]—guilty pleas.

The criminal legal system is ill-suited and ill-equipped to act as an arbiter of scientific reliability. Yet following the expansion of forensic technology without scientific scrutiny or regulatory oversight, criminal courts have become the "ground zero" of forensic reliability testing, with devastating results.[268] History has shown that criminal courts are poor gatekeepers of forensic science.[269] Jurors are poor judges of scientific reliability, and the existing rules are inadequate safeguards even for the evidence they were designed to regulate. Rather than replicate these rules with machines in mind, this Article proposes that novel approaches are needed.

IV. NOVEL APPROACHES TO NOVEL ISSUES

There is no easy solution to the issues posed by automated junk science. Every facet of the criminal legal system has enabled the pervasive role of junk science in criminal prosecutions and punishment. Forensic developers have evaded scientific scrutiny, law enforcement and prosecutors have readily adopted forensic tools without scientific validation, defense challenges have been met with legal and practical barriers, courts have shirked meaningful gatekeeping responsibilities in favor of legal precedent, and jurors have assumed the legitimacy of forensic evidence at trial. When flaws in forensic evidence were first revealed, actors in the legal system had an opportunity to reckon with the reality of junk science in criminal courts and do something about it. Yet junk science remains pervasive in criminal courts. Efforts to address the problem of junk science in criminal courts have taken the form of scientific committee inquiries,[270] implementation of more rigorous admissibility standards,[271] limitations on forensic expert testimony,[272] publication of aspirational forensic

266. Somil Trivedi, *Coercive Plea Bargaining Has Poisoned the Criminal Justice System. It's Time to Suck the Venom Out.*, ACLU: NEWS & COMMENT. (Jan. 13, 2020), https://www.aclu.org/news/criminal-law-reform/coercive-plea-bargaining-has-poisoned-the-criminal-justice-system-its-time-to-suck-the-venom-out (discussing the coercive tools impacting a decision to plead guilty, regardless of actual guilt).

267. *See* Boykin v. Alabama, 395 U.S. 238, 242 (1969) (establishing requirement that a guilty plea be knowing, voluntary, and intelligent); Bordenkircher v. Hayes, 434 U.S. 357, 364 (1978) (finding that the prosecutor's threat to re-indict Hayes on a charge carrying a mandatory life term if he did not plead guilty was not coercion).

268. *See* THE INNOCENCE PROJECT, *supra* note 32.

269. Balko, *supra* note 38; *see generally Post-PCAST Court Decisions Assessing the Admissibility of Forensic Science Evidence*, NAT'L CTR. ON FORENSICS: A PROGRAM OF THE NAT'L INST. OF JUST. (June 22, 2022), https://nij.ojp.gov/program/national-center-forensics/post-pcast-court-decisions-assessing-admissibility-forensic-science-evidence (for a database of court decisions following the PCAST REPORT).

270. *See, e.g.*, PCAST REPORT, *supra* note 33; NRC REPORT, *supra* note 25.

271. *E.g.*, Daubert v. Merrell Dow Pharms., Inc., 509 U.S. 579, 589 (1993).

272. *See, e.g.*, Scurich et al., *supra* note 171.

standards,[273] direct litigation,[274] proposed legislation,[275] investigative journalism,[276] and critical scholarship.[277]

Faced with emerging forensic automation, the legal community is newly presented with opportunities to confront both novel and enduring issues. This Part suggests reforms to address risks posed by automation: first, reforms to more appropriately distribute forensic decision-making power, and second, reforms to legal rules to account for the nature of automation. These reforms, which can be fueled by extra-system organizing, offer tools to mitigate the risks of automated forensic evidence but fall short of the only real solution: the removal of junk science from the criminal legal system.

A. REDISTRIBUTE FORENSIC DECISION-MAKING POWER

Criminal courts are called upon to "think like scientists"[278] when considering forensic reliability. Actual scientists[279] have revealed that courts have, understandably,[280] failed to do so effectively. To pursue better outcomes—only allowing the use of independently scientifically validated forensic methods in criminal casework—we must redistribute power over forensic decision-making to entities that lack adversarial incentives and possess subject-matter expertise. Decisions about forensic transparency, discovery, and protocols should be redistributed to legislators. Decisions about scientific reliability as a prerequisite to admissibility should be redistributed to scientists.

1. Legislative Oversight of Automated Evidence

Legislation can address (but not solve) some of the issues raised by automated junk science. Discovery decisions can be redistributed from criminal courts to legislators. Decisions about requirements and standards for forensic examiners, labs, and practices can be redistributed from law enforcement agencies to legislators. Central to legislative agendas must be transparency. Pursuit of algorithmic transparency in individual cases through defense discovery demands has been unsuccessful or limited in scope,[281] inconsistent,[282]

273. *See, e.g.*, NAT'L INST. OF STANDARDS & TECH., *supra* note 56.

274. *See, e.g.*, State v. Ghigliotty, 232 A.3d 468, 471 (N.J. Super. Ct. App. Div. 2020); State v. Pickett, 246 A.3d 279, 283 (N.J. Super. Ct. App. Div. 2021); State v. Loomis, 881 N.W.2d 749, 753 (Wis. 2016); People v. Wakefield, 38 N.Y.3d 367, 371 (2022).

275. *See, e.g.*, Press Release, Mark Takano, House of Representatives, Black Box Algorithms' Use in Criminal Justice System Tackled by Bill Reintroduced by Reps. Takano and Evans (Feb. 15, 2024).

276. *See, e.g.*, Balko, *supra* note 38; Segura & Smith, *supra* note 37.

277. *See, e.g.*, Lau, *supra* note 18; Sinha, *supra* note 25; Gutierrez, *supra* note 241 (arguing that the low error rates cited by firearms examiners are artificially deflated due to the use of overly simplistic and unchallenging comparisons in validation studies, and that more rigorous studies reveal misidentification rates as high as 39.6%).

278. *See* Jasanoff, *supra* note 160, at S50.

279. *See supra* Subpart.III.A.

280. This is understandable because judges lack scientific education, training, and experience.

281. *See supra* text accompanying note 203.

282. *See supra* text accompanying note 202.

and lacking in broader jurisdictional application, illustrating the inefficiency and inconsistency of addressing transparency concerns through case-by-case litigation. Legislative and ongoing regulatory oversight is needed.

Currently, forensic protocols and standards are merely optional. Firearm experts are not required to possess any certification or educational background, or follow any centralized standards or methodology in order to work or testify in criminal cases.[283] Laboratories are not required to receive any certification or oversight to operate.[284] Forensic standards and protocols published by discipline-specific committees funded by NIST are optional.[285] If a local law enforcement agency is concerned with ensuring the reliability of its forensic methods and tools, it can choose to adopt NIST standards. If not, it need not.[286] Legislation could change that.

Legislation could also address a number of the issues raised in Part III by passing laws that, at least:

- Prohibit or limit the application of trade secrets rights to forensic tools used in prosecution or investigation, or introduce novel avenues of protecting intellectual property rights owners' interests in trade secrets that are disclosed in connection with criminal prosecution or investigation.
- Require public disclosure of forensic companies' algorithm source codes before they are implemented in case work.
- Implement mandatory standards for forensic practitioners, laboratories, and methodologies, informed by the broader scientific community (rather than merely forensic practitioners themselves).
- Require rigorous, independent scientific validation standards for new forensic tools and methods, informed by the broader scientific community (rather than merely forensic practitioners themselves).

Enacting such laws is no easy feat. To date, legislative efforts have been unsuccessful. In 2024, Congressmen Mark Takano (D-Calif.) and Dwight Evans

283. *See* NRC REPORT, *supra* note 25, at 6 ("Indeed, most jurisdictions do not require forensic practitioners to be certified, and most forensic science disciplines have no mandatory certification programs."). While individual agencies may choose to implement qualifications for hiring, that is entirely discretionary. And while AFTE publishes its "Theory of Identification," *see supra* Subpart.I.A, it is also not binding on examiners.

284. *See* NRC REPORT, *supra* note 25, at 6 ("Moreover, accreditation of crime laboratories is not required in most jurisdictions.").

285. NIST is explicitly non-regulatory. While NIST develops standards for various sectors, including the use of technology in criminal case work, these standards are entirely optional. NIST has no power to require that forensic practitioners apply their standards, nor any power to control what or how technology is used by police and prosecutors. *See supra* text accompanying note 158.

286. For many years in the jurisdiction where *Ghigliotty* was litigated, the local prosecutor's office called the same firearm toolmark expert to testify in nearly every case involving a gun. That expert was not certified by any forensic entity, nor did he follow NIST or otherwise issued protocols. Instead, he wrote his own manual. This never seemed to concern the prosecutors or judges. The extent to which it troubled jurors is unknown because attorneys are forbidden from speaking with jurors after trial, but the countless convictions that rested, at least in part, on this expert's testimony suggest they were also not particularly persuaded by defense cross-examination or argument on these failures.

(D-Penn) introduced in the United States House of Representatives the Justice in Forensic Algorithms Act of 2024,[287] which would:

> Prohibit[] the use of trade secrets privileges to allow the defense access to source code and other information about software used to process, analyze, and interpret evidence in criminal proceedings;
>
> Direct[] the National Institute of Standards and Technology to establish both Computational Forensic Algorithm Testing Standards and a Computational Forensic Algorithm Testing Program; and require[] federal law enforcement to comply with standards and testing requirements in their use of forensic algorithms.[288]

There has been no action on the bill.[289] Other federal legislative efforts have focused on regulating algorithmic tools outside the criminal legal system, also without success.[290]

Legislative efforts might see more success at the local level. While not governing *forensic* algorithms, in 2017, the New York City Council unanimously passed the algorithmic accountability bill,[291] which focused on algorithmic systems used by the city to make decisions regarding "the allocation of everything from police officers and firehouses to public housing and food stamps."[292] Troubled by bias in algorithmic machine systems, Council Member James Vacca sponsored the bill and declared at a hearing on the bill: "If we're going to be governed by machines and algorithms and data, well, they better be transparent."[293] The bill created a merely advisory Automated Decision Systems ("ADS") Task Force, and a variety of procedures and criteria governing the use of algorithms in decision-making processes.[294] But none of the recommendations or proposed procedures address the use of algorithms by law enforcement or in criminal courts.

---

287. Justice in Forensic Algorithms Act of 2024, H.R. 7394, 118th Cong. § 2(a).

288. Takano, *supra* note 275.

289. Justice in Forensic Algorithms Act of 2024, H.R. 7394, 118th Cong. (referred to the Committee on the Judiciary and the Committee on Science, Space, and Technology on February 15, 2025).

290. Algorithmic Accountability Act of 2023, H.R. 5628, 118th Cong. § 8(a) (2023–2024) (proposing a law that would require companies to assess the impacts of AI systems, create new transparency about when and how such systems are used, and empower consumers to make informed choices when they interact with AI systems, and establish a Bureau of Technology within the Federal Trade Commission to enforce and regulate).

291. N.Y.C. Council, Int. No. 1696-2017 (N.Y. 2018).

292. Julia Powles, *New York City's Bold, Flawed Attempt to Make Algorithms Accountable*, NEW YORKER (Dec. 20, 2017), https://www.newyorker.com/tech/annals-of-technology/new-york-citys-bold-flawed-attempt-to-make-algorithms-accountable.

293. *Id.*

294. The Task Force was required to write a set of recommendations to address concerns about transparency and bias, including, among other recommendations, creation of: a procedure for impacted individuals to request information on decisions involving automated decision systems, a procedure for the City to determine any disproportionate impact on protected categories of persons, and a process for publicly disclosing information about agency systems. N.Y.C. AUTOMATED DECISION SYS. TASK FORCE, NEW YORK CITY AUTOMATED DECISION SYSTEM TASK FORCE REPORT 17 (2019), https://www.nyc.gov/assets/adstaskforce/downloads/pdf/ADS-Report-11192019.pdf.

Legislative efforts to address the risks of algorithmic tools have been unsuccessful, focused outside the criminal legal system, or both. To date, no laws compel defense access to source codes used by algorithmic policing tools, and none mandate standards, disclosures, or oversight of forensic technology and experts. Such legislation would be a step toward better protection of the criminally accused and evidentiary integrity.

Absent legislative action, government officials could incentivize the same results by conditioning government funding of law enforcement and prosecutors' offices on their demonstrated compliance with existing recommended standards and certification requirements.

2. Scientific Oversight of Automated Evidence

Redistributing legal decisions about a forensic method's reliability from courts to scientists can help address (but not solve) the issues raised by automated junk science. When finally consulted, scientists were alarmed by the state of forensic evidence in criminal courts.[295] Yet scientists bear little influence over forensic decisionmaking. Admissibility decisions are left to judges, who can choose to credit or discredit scientific input, and whose decisions are given great deference on review.[296]

Shifting some decisionmaking power away from courts and into the hands of the scientific community could theoretically offer more meaningful scientific screening. One way to achieve this might be to shift reliability analyses under *Daubert* to the purview of scientists rather than judges. This could be accomplished by creating committees of independent experts with diverse, relevant expertise, with the authority to limit what forensic evidence can proceed to legal screening in criminal courts. A scientific consensus that a forensic method or tool is empirically demonstrated to be foundationally valid would be a prerequisite for legal admissibility. The judge's role would be limited to resolving the *legal*, not scientific, issues implicated by forensic evidence in a criminal trial: Is the evidence impermissible hearsay? Is it more prejudicial than probative? Does admission violate the Confrontation Clause? A second option is the incorporation of scientists in admissibility analyses by establishing specialty courts to deal with issues of forensic evidence, including a formal role for scientists on court panels.

While offering the hope of injecting some actual scientific screening into admissibility decisions, these proposals raise questions of practicality (how can we implement such a massive restructuring of authority?) and unintended risks (would this, again, falsely legitimize junk science by simply cloaking it in layers of scientific review?). In theory, mandatory scientific screening should accomplish the ultimate solution of abolishing junk science. But there are

295. *See supra* Subpart.I.A.2.

296. *See* General Elec. Co. v. Joiner, 522 U.S. 136, 142–43 (1997) (explaining admissibility decisions by trial judges are reviewable only for abuse of discretion).

reasons to fear, perhaps even expect, different results. Meaningful progress can only be achieved if committees or panels are independent, qualified, and well-intentioned. These are not always features of criminal legal system actors and governmentally-formed entities.

Redistribution of power to scientists could also be accomplished to some degree simply by ethical prosecutors abandoning the historical "trial and error" approach to forensics—implementing forensic methods at their inception and only reconsidering their use when conclusively demonstrated to be unreliable—and replacing it with caution and scrutiny. Use of automated forensic tools should be premised on proof of independent, scientific validation. Prosecutors and law enforcement agencies bear a responsibility to demand exacting scientific validation before purchasing and implementing these tools. Courts bear a responsibility in demanding exacting scientific validation before admitting the resulting evidence. Such demands would create incentives for forensic manufacturers to prioritize independent, transparent scientific review.

Though legitimizing, reforms redistributing scientific screening to independent scientists might at least reduce admission of flawed forensic evidence and decouple screening from law enforcement interests.

B. REFORM LEGAL RULES TO ACCOUNT FOR AUTOMATION

Another approach offers reforms within the existing system. These reforms—new or modified rules contemplating the novel issues raised by automation—provide tools for litigants and courts to reduce the harm of junk science in criminal case outcomes.

1. Discovery Rules Tailored to Automation

Discovery rules directly addressing the nature of automated evidence are needed. Constitutional imperatives[297] demand allowing discovery of algorithm source codes. Exact language for such a discovery rule should be drafted in collaboration with independent scientific experts, but the rules should unambiguously require disclosure of, at least, source codes, machine maintenance records, all error rate and validation studies conducted on the tool (whether or not published), reports documenting every "filter" and setting applied in the case, and all images created during the examination (whether used in evidence or not). Tailored discovery rules can conclusively arm the accused with discovery necessary to defend against incriminating forensic evidence.

2. Admissibility Rules Tailored to Automation

Absent systemic change to admissibility decision-making authority, new rules are needed that incorporate explicit legal consideration of reliability issues unique to automation—namely, sources of algorithmic, machine, and human

297. *See supra* Subpart.III.B.2.

error. New admissibility rules should clearly define the "relevant scientific community" considered by *Daubert* and *Frye*. These new rules should explicitly include not only firearm examiners, as often narrowly interpreted by prosecutors and judges, but also historically excluded experts who can provide courts with the necessary expertise and insights, such as computer programmers, statisticians, academics, and social scientists. Federal Rule 707[298] offers an imperfect and incomplete reform but demonstrates openness of lawmakers to adapt admissibility rules to the realities of automation. The logical next step is to amend or supplement Rule 707 with algorithm-specific criteria, requiring judges to consider sources of algorithm and machine error, hear from a broader array of experts, and undergo some independent education before making final admissibility determinations.

3. Jury Instructions Tailored to Automation

Just as model jury instructions provide factors to consider in determining whether to believe a human witness's testimony, and how much weight to give it,[299] so too could instructions be drafted to guide jurors' assessment of machine aided or created evidence. While algorithms and machines may not be capable of intentional deceit like human witnesses, they are certainly capable of aiding or uttering inaccurate or misleading outputs. Judges could instruct jurors of indicators of credibility for machine witnesses. Learning the limitations of automated evidence from the judge, rather than adversarial litigants, can lend credibility and provide some control for variation in attorney competency and strategy. Even if a defense attorney fails to point out reliability concerns with automated evidence, jurors will hear them from the judge.

The use of jury instructions to address issues in forensic evidence is not unprecedented. In 2011, the New Jersey Supreme Court issued a landmark ruling, *State v. Henderson*, that incorporated extensive scientific critique of eyewitness identifications and mandated comprehensive model jury instructions educating jurors and guiding their credibility assessments of eyewitnesses.[300] *Henderson*-like jury instructions could be implemented for automation, enumerating the sources of potential algorithm, machine, and human error unique to algorithmic tools.

Admittedly, the impact of such jury instructions is unclear.[301] Do they truly educate and inform juror decisionmaking? Or do they empower judges to shift

298. *See supra* Subpart.III.C.1.

299. Sinha, *supra* note 25, at 932.

300. 27 A.3d 872, 877–88 (N.J. 2011). *Henderson* also mandated a more rigorous admissibility process for eyewitness identifications. *Id.*

301. The *Henderson* decision led to more extensive and meaningful pretrial hearings for eyewitness testimony in New Jersey, but the experience of litigants shows that eyewitness identifications are still usually admitted. The impact of the jury instructions on jury deliberations is unknown; attorneys are not permitted to talk to jurors after trial.

accountability to jurors, excusing lax gatekeeping? At the very least, model instructions would allow for more informed jury deliberations.

Each of these proposed reforms offers an opportunity to alleviate harms for those most impacted—namely, the criminally accused and over-policed communities—but they also risk legitimizing the use of forensic tools.[302] Recognizing this tension, non-reformist reforms "aim to undermine the prevailing political, economic, social order, construct an essentially different one, and build democratic power toward emancipatory horizons. They seek to redistribute power and reconstitute who governs and how."[303] In the context of automated forensic evidence, non-reformist reforms can challenge the existing power structures that have allowed junk science to persist in criminal courts by fundamentally altering who has authority over forensic decision-making and how that authority is exercised. The extent to which the reforms offered are actually non-reformist will depend on their implementation. While risks of greater entrenchment and false legitimization of junk science remain, given the cost of human inaction, [304] protective action is compelled.

C. ORGANIZE TO DEMAND ACCOUNTABILITY

Systemic change often begins outside the legal system through community organizing and initiatives. These powerful tools, unbeholden to precedent, can lead and shape efforts to address the risks raised by automated junk science. Obscurity, secrecy, and unreliability of policing tools is of interest to the broader public, and the broader public has a role to play in demanding higher standards for forensic tools. Interest in police accountability has grown in recent years, fueled by the Black Lives Matter movement[305] and public outcry following a series of high-profile police killings of Black men and women.[306] Community organizations and journalists have since revealed the breadth of police surveillance and racial bias in police technology.[307] Journalistic scrutiny, public attention, political organizing, and public records litigation have all become effective techniques in a broader public push for transparency in police

302. Sinha, *supra* note 25, at 940 (analyzing how policing technologies entrench and exacerbate carceral harms and a framework for non-reformist reform that acknowledges this).

303. Amna A. Akbar, *Non-Reformist Reforms and Struggles over Life, Death, and Democracy*, 13 YALE L.J. 2497, 2507 (2023) (including several non-reformist reforms that are "conceived, not in terms of what is possible within the framework of a given system and administration, but in view of what should be made possible in terms of human needs and demands."); Amna A. Akbar, *Demands for a Democratic Political Economy*, 134 HARV. L. REV. F. 90, 104 (2020).

304. *See supra* note 32.

305. *About Black Lives Matter*, BLACK LIVES MATTER GLOB. NETWORK FOUND., https://blacklivesmatter.com/about ("Black Lives Matter Foundation is an abolition-centered foundation fighting institutional injustice and serving Black people globally.").

306. *High Profile Killings by Police*, POLICE BRUTALITY CTR., https://policebrutalitycenter.org/police-brutality/high-profile-killings (including the police killing of Eric Garner, Tamir Rice, and Michael Brown in 2014, Freddie Gray in 2015, Philando Castile in 2016, and Breonna Taylor and George Floyd in 2020, among many others).

307. *See, e.g.*, Angwin et al., *supra* note 139.

technology and procedure. Public scrutiny is an essential step toward needed reforms. "Transparency litigation"[308] to compel disclosure of government documents is a valuable tool to address shadow forensics and emerging automation by applying public pressure on lawmakers and judges to implement greater safeguards,[309] and to educate the public, members of which may vote on proposed reforms, serve on juries, or simply bear an interest in the practices of their local police department. Actors outside the legal system can demand transparency and reliability in forensic technology and move us incrementally closer to abolishing junk science.

D. GET RID OF JUNK SCIENCE

Junk science has no place in the prosecution and punishment of human beings. Any reform will be inadequate so long as the methods of criminal investigations and prosecutions remain as they are—scientifically unfounded. As the authors of the PCAST report concluded:

> [N]either experience, nor judgment, nor good professional practices (such as certification programs and accreditation programs, standardized protocols, proficiency testing, and codes of ethics) can substitute for actual evidence of foundational validity and reliability.[310]

The only true solution is clear: remove junk science from the toolkit of law enforcement. This solution can be grounded in abolitionist principles as a call for the elimination of forensic evidence,[311] or simply as a logical appeal to the moral, legal, and scientific imperative of evidentiary integrity.

Forensic methods and tools—mostly junk science—were developed and implemented as carceral tools.[312] Application of prison abolition principles compels abolition of forensic tools in criminal prosecutions. Admittedly, application of prison abolition principles to forensic evidence is imperfect. For one, forensic evidence can be used by the accused to defend against criminal charges or prove innocence.[313] Some forensic methods (not firearms examination) are used in non-criminal contexts. But the use of forensic evidence is fundamentally, inextricably tied to carceral purposes, implicating abolitionist concerns. What, then, does the abolition of junk science look like?

---

308. Hannah Bloch-Wehba, *Visible Policing: Technology, Transparency, and Democratic Control*, 109 CALIF. L. REV. 918, 922 (2021).

309. *See* Jonathan Manes, *Secrecy & Evasion in Police Surveillance Technology*, 34 BERKELEY TECH. L.J. 503, 512 (2019) ("Until there is a critical mass of public disclosure and public awareness, courts and legislatures generally do not publicly weigh in on the constitutional or statutory limits on the police's use of the novel technology.").

310. PCAST REPORT, *supra* note 33, at 6.

311. *See generally* Sinha, s*upra* note 25.

312. *Id.* at 886 ("Forensic methods are squarely classifiable as carceral tools; they uniquely support law enforcement activity—investigation, prosecution, and punishment . . . .").

313. *See Forensic Science*, NAT'L ASS'N OF CRIM. DEF. LAWS., https://www.nacdl.org/Landing/ForensicResources (last visited July 24, 2025) ("Evidence is the crux of every criminal case, making forensic science one of the most (if not the most) critical elements of an investigation and defense.").

Abolition of junk science can take different forms. Immediate elimination of junk science is unrealistic—though it should not be controversial to disclaim demonstrably unsound methods!—because it is so deeply entrenched in our policing and court systems. Instead, abolition of junk science might be accomplished incrementally by applying an abolitionist lens to forensic reforms.[314]

Critics of the broader abolition movement will surely lend similar critiques to this narrower abolitionist proposal—namely, that abolition of junk science is impractical and unrealistic.[315] But the greatest barrier to abolishing junk science is not the inability of the criminal legal system to disclaim it, but rather its unwillingness. Whether immediate, incremental, or incomplete, abolition of junk science will only be possible in a legal system that values accuracy over efficiency, justice over precedent, and liberty over commercial interests. Abolition of junk science is impractical only so long as our system—and those with the power to change it—continue to pursue its historical mass incarceration agenda, unrealistic only so far as we are limited by the (mis)conception of "justice" long endured.

One need not subscribe to abolitionist principles to demand forensic reliability. Whether from the perspective of abolition, scientific integrity, racial justice, safeguarding constitutional rights, police and prosecutorial accountability, public transparency, or shrinking governmental control, there are a myriad of reasons to call for the elimination of junk science in criminal courts. Simply put, only by disclaiming unscientific evidence—junk science—can the criminal legal system offer any confidence in its outcomes.

CONCLUSION

Automation is not the solution to junk science; rather, it risks obscuring and reanimating historical injustices.[316] But the new wave of forensic technology also presents the legal community with an opportunity to learn from history rather than repeat it, and to finally reckon with the stubborn persistence of junk science in criminal courts. Our legal system must confront not only its failure to meaningfully address junk science, but also its unwillingness. Until we are more concerned with reducing wrongful convictions than enabling prosecutions, and until we prioritize scientific concerns over carceral pursuits and accuracy over efficiency, we cannot claim confidence in criminal system outcomes. And until we disclaim junk sciences entirely, we cannot rely on automation to address the infirmities of human forensic methods.

314. *See* Sinha, *supra* note 25, at 938–39 (suggesting a three-part framework for reimagining forensic evidence through an abolitionist lens).

315. *Id.* at 951 (discussing critiques of a forensic evidence abolition framework).

316. *See* THE INNOCENCE PROJECT, *supra* note 32; *supra* text accompanying note 44.

***