28 May 2024

Mark Greene, Office Director
Office of Technology and Standards, National Institute of Justice
810 7th Street, NW
Washington, DC 20531

RE: IEEE-USA Input–Section 7.1(b) of E.O. 14110 (use of AI systems in criminal justice)

The use of AI systems–more specifically algorithmic determinations and automated decision making systems–in criminal justice determinations pose significant risks and harms and present distinct constitutional, privacy, electronic surveillance, civil liberties, and evidentiary concerns. Advancements in generative and predictive AI systems accelerate these risks. Specifically, the use and deployment of AI systems in surveillance and criminal justice determinations–without clear legislative safeguards and oversight–violates or sidesteps fundamental U.S. Constitutional rights), federal electronic surveillance laws and state counterparts, and evidence rules regarding relevance, reliability, privilege, and admissibility. To preserve fundamental rights, IEEE-USA recommends:

1. Strong data privacy protection.
2. Transparency and disclosure of (a) any and all AI systems used in criminal justice determinations, (b) an individual's data/evidence that has been collected by AI systems, and (c) an individual's data used in or generated by any AI system for criminal justice determinations.
3. Require that data (evidence) collected by, and data (evidence) generated by AI systems comply with the rules of evidence.
4. Require using standards such as IEEE Std. 1012 for independent testing and auditing for any AI system procured by law enforcement and any AI system used in a criminal justice determination.
5. Funding the development of standards and certifications, testing and evaluation, certification, recurring benchmarking exercises and independent studies of AI systems, while promoting AI education and removing barriers to ensuring public awareness, access, and research on the existence of, and societal impacts of AI systems.

Advancing technologies require the law to adapt, but they should not eviscerate fundamental rights. Without the steps above, the criminal justice system—cannot comply with existing laws and preserve the fundamental rights of life and liberty. Please do not hesitate to contact Erica Wissolik at (202) 360-5023 or e.wissolik@ieee.org if you have any questions.

Sincerely,

Keith Moore
IEEE-USA President

IEEE-USA | 2001 L Street, N.W., Suite 700, Washington, D.C. 20036-4928 USA
Office: +1 202 785 0017 | Fax: +1 202 785 0835 | E-mail: ieeeusa@ieee.org | Web: https://ieeeusa.org

IEEE-USA RESPONSE TO NIJ'S RFI RE: §7.1(b) of E.O. 14110

I. Overview

IEEE-USA submits the following per NIJ's RFI regarding section 7.1(b) of Executive Order 14110 (EO) addressing use of AI in the criminal justice system. The EO defines the term "AI system" as "any data system, software, hardware, application, tool, or utility that operates in whole or in part using AI." In this response, IEEE-USA's use of the term "AI system(s)" refers specifically to algorithmic determinations and automated decision making systems.

A. Use of AI Systems in Surveillance and Criminal Justice Determinations

The use of AI systems, including algorithmic determinations and automated decision making systems, in the criminal justice system poses significant constitutional and privacy concerns, particularly in making determinations with respect to guilt, sentencing, parole, release, probation, risk assessments related to determinations of suspect or criminal defendant rights, crime forecasting, predictive policing, and prison management decisions (collectively hereinafter "criminal justice determinations").

Use of AI systems in police investigation, identification, and surveillance (collectively "surveillance") and in forensic analyses pose additional distinct constitutional, privacy, electronic surveillance, civil liberties, and evidentiary concerns.

The types of AI systems used in the criminal justice system for both surveillance and criminal justice determinations are vast, including, but are not limited to, systems claiming to be able to identify people based on DNA mixtures (probabilistic genotyping systems), faces and voice (face and voice recognition systems), fingerprints (automated fingerprint identification systems), and other biometrics like retinas or gait; identify people's internal emotional states (emotional recognition systems); identify their physical appearance based on DNA (forensic DNA phenotyping); locate objects based on license plates (automated license plate recognition), audio (acoustic gunshot detection), or images (image recognition systems); and, track and profile people based on data about them (including their political opinions, religion, race, gender, income, address, and medical conditions).

B. Understanding the Nature of the Technology

AI systems, especially predictive and generative AI systems, bring forth significant challenges concerning privacy, provenance attribution, erroneous assertions (referred to in the popular press as hallucination or confabulation), and are subject to disinformation attacks. The burden has to be placed on both the manufacturer, seller, and deployer of said technology to ensure it complies

with constitutional rights, evidentiary rules, federal and state statutes (including electronic surveillance laws), and other data laws, meeting or exceeding contemporary regulatory and IP requirements. It is imperative to understand and account for trade-offs while making decisions transparently. For example, guaranteeing strict differential privacy while ensuring the fairness of a decision concerning an underrepresented group requires a trade-off. Accuracy and designer loss functions (the focus of many deep learning efforts) are often not enough. Understanding the impact of decisions is critical. Finally, there is a clear and pressing need for funding to develop independent testing, public and private auditing of said technology, and impact assessments to address these challenges.

IEEE-USA emphasizes the need for immediate action to establish robust testing and auditing protocols for AI systems used within the criminal justice system–particularly those used for surveillance and in criminal justice determinations. This urgency is particularly relevant as Americans' fundamental rights are increasingly determined by AI systems that have significant impacts on their rights. It is imperative such systems are not only reliable and trustworthy, but also demonstrably fair and equitable in ways that are accountable to the public. Further, AI developers and deployers should be assured of a leveled playing field in an environment with transparent and well-established rules. Such rules include having access to clearly defined testing and audit criteria and processes. These measures should balance societal risks and the need for innovation, while simultaneously safeguarding the confidentiality of proprietary competitive information.

C. Structure of IEEE-USA's Response

This Response first summarizes federal, state, and local law enforcement's increasing use of AI systems in surveillance and criminal justice determinations combined with federal, state, and local law enforcement's increasing access to and purchase and use of publicly and commercially available data (Section II). These AI systems include government-owned and third party vendor-supplied AI systems. Section III highlights existing constitutional rights, federal and state laws, and evidentiary rules and concepts that are violated or eroded by this combination of AI systems plus commercially available data, which frequently contains data that the law otherwise would protect from government collection without a warrant, as well as the risks and harms this combination poses in criminal justice determinations. Section IV sets forth IEEE-USA's recommendations for federal legislation to address these risks.

II. Proliferation of AI Systems in Criminal Justice and its Combination with Commercially Available Data: A Case Study

A. The Criminal Justice System, the Consumer Data Market, and AI Systems

Numerous government agencies, including the FBI, Department of Defense, National Security Agency, Treasury Department, Defense Intelligence Agency, Navy and Coast Guard, have purchased vast amounts of U.S. citizens' personal information from commercial data brokers. The revelation was published in a partially declassified, internal Office of the Director of National Intelligence Report released on June 9, 2023.

The report captures both how widespread government purchases of commercially available information are and how haphazard government practices around the use of the information are. The purchases are so pervasive and agencies' practices so poorly documented that the Office of the Director of National Intelligence cannot even fully determine how much and what types of information agencies are purchasing, and what the various agencies are doing with the data. Law enforcement agencies also contract with private companies, like Fog Reveal and Kochava, to obtain location information and other data for which the law would otherwise require officers to obtain a warrant.

The report shows the large-scale and invasive nature of the consumer data market, which collects, aggregates, buys, analyzes, and sells *publicly available information* and *commercially available information*. The distinction between the two types of data is significant from a legal perspective: publicly available information is information that is already in the public domain, whereas commercially available information is personal information collected from a dizzying array of sources by commercial data brokers that aggregate and analyze it, then make it available for purchase by others, including governments.

The sources and types of commercially available information (data) are mind-bogglingly vast. They include public records and other publicly available information. But far more information comes from the nearly ubiquitous internet-connected devices in people's lives, like cellphones, smart home systems, cars and fitness trackers. These all harness data from sophisticated, embedded sensors, cameras and microphones. Sources also include data from apps, online activity, texts and emails, and even health care provider websites.

Types of data include location, gender and sexual orientation, religious and political views and affiliations, weight and blood pressure, speech patterns, emotional states, behavioral information about myriad activities, shopping patterns and family and friends.Some of that information is private, confidential or otherwise legally protected.

This data provides companies and governments a window into the "Internet of Behaviors," a combination of data collection and analysis aimed at understanding and predicting people's behavior. It pulls together a wide range of data, including location and activities, and uses scientific and technological approaches, including psychology and machine learning, to analyze that data. The Internet of Behaviors provides a map of what each person has done, is doing and is expected to do, and provides a means to influence a person's behavior.

III. Use of AI Systems in Criminal Justice Circumvents Constitutional Rights, Statutes, and Evidentiary Rules

A. Constitutional and Statutory Rights in Criminal Justice and Restrictions on Surveillance

Government use of electronic surveillance tools is extensively regulated by federal and state laws. The U.S. Supreme Court has ruled that the Constitution's Fourth Amendment, which prohibits unreasonable searches and seizures, requires a warrant for a wide range of digital searches. These include wiretapping or intercepting a person's calls, texts or emails; using GPS or cellular location information to track a person; or searching a person's cellphone.

The jurisprudence surrounding the First, Fourth, Fifth, and Sixth Amendments to the Constitution, rights to free speech, liberty, and privacy, and law enforcement has a long history and a full overview is beyond the scope of this article. So too is the U.S. domestic electronic surveillance scheme's jurisprudence. But to understand the disconnect that currently exists between those constitutional and statutory rights and the use of AI systems in criminal justice regulation of the commercial data market (data privacy), some background is necessary.

The Fourth Amendment prohibits unreasonable searches and seizures, but it applies only to government search and seizure. In its framework of jurisprudence addressing enhanced forms of surveillance, the Supreme Court has ruled that the government's warrantless uses of wiretaps,[1] dog sniffing,[2] thermal imaging,[3] attachment and use of a physical GPS device,[4] and obtaining cell site location information (CSLI) for tracking purposes[5] are all unlawful violations of the Fourth Amendment.

The Electronic Communications Privacy Act (ECPA) regulates the interception of wire, oral, and electronic communications by government and private actors.[6] Through ECPA, Congress has sought to safeguard the privacy of wire, oral, and electronic communications and, in particular, the privacy of innocent persons. ECPA forbids the interception of wire, oral or electronic communications by private persons unless the communication is intercepted by, or with the consent of, a participant, and significantly restricts the authority of law enforcement officials to intercept such communications. ECPA and the many state laws that mirror it regulate when and how domestic law enforcement and private entities can "wiretap," i.e., intercept a person's

[1] *Katz v. United States*, 389 U.S. 347 (1967).
[2] *Florida v. Jardines*, 569 U.S. 1, 11 (2013).
[3] *Kyllo v. United States*, 533 U.S. 27, 40 (2001).
[4] *United States v. Jones*, 565 U.S. 400, 412–13 (2012).
[5] *Carpenter v. United States*, 138 S. Ct. 2206, 2223 (2018).
[6] For an in-depth analysis of electronic surveillance law in the U.S., see *Wiretapping & Eavesdropping: Surveillance in the Internet Age, 3rd. Ed.*, McKenna, Anne Toomey and Fishman, Clifford S. (Thomson | Reuters), including § 1:15 et seq Electronic Communications Privacy Act (ECPA).

communications, or track a person's location. In short, ECPA restricts what, when, and how data can be collected.

Coupled with Fourth Amendment protections, ECPA generally requires law enforcement agencies to get a warrant based on probable cause to intercept someone's communications or track someone's location using GPS and cell site location information. Also, ECPA permits an officer to get a warrant only when the officer is investigating certain crimes, so the law limits its own authority to permit surveillance of only serious crimes. Violation of ECPA is a crime, but ECPA permits interceptions and electronic surveillance when a person consents to that surveillance.

B. Electronic Surveillance v. Data Privacy: How AI Systems and Commercial Data Circumvent Existing Law

Electronic surveillance law protections and *data privacy* mean two very different things in the U.S. There are robust federal electronic surveillance laws governing domestic surveillance, but there is no federal data privacy law. This creates an enormous loophole, and government use of AI systems in surveillance (including investigation and identification) take advantage of and widen that loophole. Thus, even though our electronic surveillance laws and the Supreme Court have specifically forbidden warrantless government surveillance and tracking, because the U.S. lacks a comprehensive federal data privacy law, the ability to track most persons at all times may be permissible.

With little in the way of federal data privacy laws, once someone clicks "I agree" on a pop-up box, there are few limitations on private entities' collection, use and aggregation of user data, including location data. This distinction between data privacy and electronic surveillance law protections creates the framework that underpins the massive U.S. data sharing market.

Advanced predictive and generative AI systems rely upon, use, and provide data that would otherwise be protected from government collection and use by the Fourth Amendment and ECPA. Complying with these laws takes time and money, plus electronic surveillance law restricts what, when and how data can be collected. Commercially available information is cheaper to obtain, provides far richer data and analysis, and is subject to little oversight or restriction compared to when the same data is collected directly by the government.

Advancing AI tools are enabling a new phase in data analysis. Generative AI's ability to process vast amounts of data is reshaping what can be done with and learned from mobile data in ways that invade privacy. This includes inferring and disclosing sensitive or otherwise legally protected information, like medical records and images. It also makes it possible to manipulate individual and group behavior, inducing decisions in favor of the specific users of the AI tool.

C. Use of AI Systems in Criminal Justice Determinations

Criminal justice determinations made utilizing AI systems pose Fifth and Sixth Amendment concerns and conflict with the concepts of evidence law that underpin the U.S. legal system generally, including privilege, and the Federal Rules of Evidence specifically.

The Fifth Amendment protects individuals from forced self-incrimination, among other rights. Government use of AI systems that use or analyze an individual's commercially available data may be compelling that individual to provide evidence against themselves.

By the very nature of how the technology functions, the use of AI Systems in criminal justice determinations also may circumvent multiple Sixth Amendment rights, including the accused's rights to "an impartial jury" from the locale where the crime was committed, the right "to be informed of the nature and cause of the accusation," and the right "to be confronted with the witnesses against him." When investigation, identification, and other criminal justice determinations are made by AI systems relying on unknown or non-transparent inputs, outputs, and data sets, the Sixth Amendment's protections are eviscerated.

The Federal Rules of Evidence (Fed. R. Evid.) govern the discovery, disclosure, and admissibility of evidence in civil and criminal litigation. The ability of AI systems to make criminal justice determinations undermines the role, responsibility, and duty of the judge as gatekeeper of evidence to make such determinations and to adhere to the intent and purpose and Fed. R. Evid. (fair and speedy trials with just determinations based upon relevant and reliable evidence)

Relevance is the cornerstone of the law of evidence, and the concept of relevance underlies the classic definitions of "evidence."[7] Fed. R. Evid. 401 dictates two necessary steps to assess whether evidence is relevant: does the item of evidence have (1) "a tendency to make a fact more or less probable than it would be without the evidence" and is that fact (2) "of consequence in determining the action." Thus, establishing relevance requires examining the relationship between the item of evidence being offered and the "fact of consequence." The Rule summarizes this relationship as a "tendency to make the existence" of the fact to be proved "more probable or less probable."

Fed. R. Evid. 401's definition and "test" confirm that "relevancy" does not exist in a vacuum. It is necessary to examine an item of evidence considering the issues in the case to determine whether that evidence meets Fed. R. Evid. 401's relevancy standards. In other words, the evidence must be assessed against the elements of the cause of action, crime, or defenses at issue in the trial, but that assessment should be flexible and open-ended as the case and the issues presented evolve and progress in the litigation process.

[7] See, Jones on Evidence § 1:4 (7th ed.), Fishman & McKenna (West | Thomson).

Expert opinion testimony plays a crucial and potentially determinative role in many civil and criminal cases, and experts are increasingly called upon to explain to jurors an AI system's analysis of or determination about evidence. When a party seeks to offer expert testimony, the trial judge must make several initial determinations, pursuant to Fed. R. Evid. 104(a). The trial judge as "gatekeeper" must first determine whether the witness is qualified to give the testimony in question. The trial judge must next determine that this testimony will help the trier of fact to understand the evidence or determine a fact in issue. Finally, the trial judge must determine that this testimony is sufficiently trustworthy, which requires that the testimony is the product of reliable principles and methods that the expert reliably applied to the specific facts of the case.

More specifically, Fed. R. Evid. 702 requires that any testimony by experts, such as those using and interpreting AI system's results for the courts, is (a) based on the "the expert's scientific, technical, or other specialized knowledge," (b) "based on sufficient facts or data," (c) "the product of reliable principles and methods;" and (d) "reflects a reliable application of the principles and methods to the facts of the case." The question of expert qualification is inextricably connected with the trial court's determination of whether the expert testimony is reliable.[8]

Courts cannot properly fulfill their gatekeeping role to determine the reliability, and therefore, admissibility, of an expert's testimony under Fed. R. Evid. 702 without a proper understanding of the AI systems, the provenance (and admissibility) of the underlying data used in the AI system, results of independent testing and auditing adhering to proper standards and certifications, introduced by an expert with the scientific and technical knowledge to understand the interpret the AI system.

IV. Federal Legislation Recommendations

To preserve fundamental rights to life and liberty, IEEE-USA recommends federal legislation that promotes:

- Strong data privacy protection.
- Transparency and disclosure of (a) any and all AI systems used in criminal justice determinations, (b) an individual's data/evidence that has been collected by AI systems, and (c) an individual's data used in or generated by any AI system for criminal justice determinations.
- Require that data (evidence) collected by, and data (evidence) generated by AI systems comply with the rules of evidence.

[8] § 43:4. Expert qualification under Fed. R. Evid. 702, 6 Jones on Evidence § 43:4 (7th ed.) Fishman & McKenna (West | Thomson).

- Require using standards such as IEEE Std. 1012 for independent testing and auditing for any AI system procured by law enforcement and any AI system used in a criminal justice determination.
- Funding the development of standards and certifications, testing and evaluation, certification, recurring benchmarking exercises and independent studies of AI systems, while promoting AI education and removing barriers to ensuring public awareness, access, and research on the existence of, and societal impacts of AI systems.

The following sections address each of these recommendations in turn.

A. Strong data privacy protection.

Despite decades of increasingly sophisticated and invasive commercial data aggregation, Congress has not passed a federal data privacy law. The lack of federal regulation around data creates a loophole for government agencies to evade electronic surveillance law. It also allows agencies to amass enormous databases that AI systems learn from and use in often unrestricted ways. The resulting erosion of privacy has been a concern for more than a decade. AI's ubiquitous presence in society has challenged our ability to protect privacy and ensure equity and justice. To ensure equity and justice in the criminal justice system requires updating, harmonizing, and streamlining federal laws, policies, and guidelines relating to privacy as follows:

1. Enact clear and comprehensive privacy and data protection law(s) at the federal level.

Equitable AI practices require a clear legislative framework for data ownership, confidentiality of data, and rights of access to data used in and by AI systems--essential to protecting privacy and autonomy. Moreover, the absence of a comprehensive data protection law at the federal level in the U.S. is a missed opportunity for the U.S. to globally shape and address data rights, practices, and privacy. The current patchwork of federal and state laws lacks coherence and is insufficient.[9]

[9] Our current federal and state patchwork of data laws lends to confusion and inefficiencies and precludes the U.S. from shaping private and government sector data practices on a national and international level. At the federal level, there is a sectoral-based approach to data regulation by both public and private sectors (e.g., health – Health Insurance Portability and Accountability Act; and financial – Gramm-Leach-Bliley Act). Given current data collection practices (communication platforms, apps, and devices routinely collect health, financial, and biometric data, and PII), the existing sectoral approach leaves vast swaths of individuals' intimate data unprotected in the current federal legislative scheme. At the state level, all 50 states have passed varying forms of data breach laws and a myriad of states have enacted comprehensive data regulation and biometric laws, such as the California Consumer Privacy Act (CCPA) and the Illinois Biometric Information Privacy Act. California, via CCPA, and the European Union's General Data Protection Regulation Act (GDPR) both provide legislative frameworks that have altered private sector data practices on a global scale.

Internet and other communication-related platforms, apps, and devices routinely collect or infer health, financial, and biometric information without user knowledge, control, or consent. The U.S. sectoral approach to data regulation leaves vast swaths of individuals' intimate data unprotected and fails to provide a clear framework of permissible operation for AI systems and their operators, leading to inefficiencies and confusion, and violations of Constitutional and statutory rights. Comprehensive data regulation through legislative action should incorporate principles like Fair Information Practice Principles (FIPPs) that:

- Establish data collection and data use limitations, data quality standards, and security safeguards.
- Require clearer notice of data collection practices with truly effective opportunities to consent (or not) to such data collection.
- Mandate transparency and user control in use of individual data. Consumers are often unaware of how their data is collected and used; long, complex Terms of Use and privacy policies obscure actual data practices. Mandate that users have the right to access, review, store, and delete personal user data, including behavioral data used for tracking and AI recommendation systems, and require an option to opt-out of tracking.

2. Create baseline standards for platform governance.

AI systems are ubiquitous, and access to and use of online platforms is a requirement to be an effective citizen of the modern world (education, taxes, banking - all require online participation with platforms, devices, and apps that operate with and rely upon AI systems). To protect both domestic and national security interests and the constitutional rights (speech and privacy) of users, baseline standards should be created for: verification procedures for account creation; when accounts can or should be removed or deactivated for a period of time; and when content can or should be removed or labeled with warnings.

B. Transparency and disclosure of (a) any and all AI systems used in criminal justice determinations, (b) an individual's data/evidence that has been collected by AI systems, and (c) an individual's data used in or generated by any AI system for criminal justice determinations.

1. Disclose information to ensure individual and public understanding of AI systems' capabilities and limitations.

Responsible development and governance of AI systems requires ensuring transparency in how this technology makes decisions, as well as public understanding of its capabilities and limitations. To achieve this, the U.S. should develop public outreach strategies that inform about AI, improve broad and inclusive participation in its design and regulation, and develop

appropriate levels of trust in the technology. In particular, the public must understand and be aware of the following:

- When particular AI systems are in use;
- AI systems' competency and the extent to which they might produce disparate impacts;
- Whether they are safe and secure, and how this is evaluated;
- Their legality and legal accountability;
- Their impacts on privacy;
- Whether they might constrain the autonomy of users or other affected individuals; and,
- Their potential ethical and societal impacts.

2. Disclose information to understand the provenance of inputs and outputs of AI systems throughout their supply chain and life cycle.

AI tools such as large language models (LLMs) have the potential to result in logical and mathematical inaccuracies in text, incorrect or impossible images, and unacceptable or hurtful biases. Understanding these unique threats requires thoughtful guardrails to be put in place to prevent risk across the model building lifecycle (pre and post deployment). This raises the need to examine how data is obtained and processed at the multiple phases of collection.

3. Disclose information to enable discovery and mitigation of disparate impacts.

When AI systems are developed and deployed, objectives of accuracy and lack of algorithmic bias towards marginalized or vulnerable groups can conflict, resulting in disparate impacts and lack of public confidence. To mitigate, objectives must be balanced by means that require clarity, transparency, and protection of all stakeholders.

- Establish and mandate metrics and standards. AI systems and their operators must comply with standards for fairness, privacy, safety, and security.
- Establish transparency mechanisms for stakeholders. For example, require third-party access to data in standardized, machine-readable format.
- Create research investments on how the use of algorithms may disparately impact or disadvantage certain individuals and groups.

4. Disclose information to enable redress.

When AI systems make life-impacting decisions, preserving privacy, equity, and justice requires that individuals be informed about, and permitted to, question decisions and have access to systems that enable redress.

- Define pathways for all stakeholders to report problems, question results, provide additional information relevant to automated decision making, and receive redress when they are harmed.
- Define pathways for individuals to review, verify and question input data about them as individuals.
- Require human teams be tasked to investigate errors with clear pathways for stakeholders to communicate with teams and require timely response.
- Require systems to produce explanations of their output that can be examined by human decision makers and other stakeholders.
- Provide clear statutory culpability and means of civil redress for entities in the AI supply chain responsible for harm to individuals, groups, or the environment.

5. Disclose information necessary to research the existence, fairness, safety, security, privacy, and ethical and societal impacts of AI systems.

Increasingly, AI systems directly impact human life, individual rights and societal well-being and, like other systems that do so, must be evaluated throughout their lifecycle, i.e., design, implementation, and deployment. When AI systems are deployed in critical applications such as the criminal justice system:

- Governments should: (i) publicly identify and disclose the AI systems used by the government; (ii) require and publicly disclose a methodological validation study that establishes the value of using new AI systems in place of existing practices prior to deploying AI systems; (iii) adopt clear procedures relating to the collection, usage, storage, and sharing of personal information in the context of developing, using, and validating a given AI systems in a privacy-preserving manner; and (iv) prevent intellectual property, confidentiality claims, lack of funding, or lack of an designated independent body within government to monitor compliance from impeding duly limited independent validation and verification and publicly disclosed review of the fairness, safety, security, privacy, and ethical and societal impacts of AI systems. AI systems ought to be submitted to the organization performing validation and verification thereof, and the organization using related private intellectual property or proprietary data in its evaluation must adopt rules to protect such private rights from misappropriation.
- Users and the public should be allowed to: (i) request and receive an explanation of how a government determination using AI systems was reached; (ii) determine whether the AI systems used in government decision-making disproportionately impacts a protected class; and (iii) rectify, challenge, or complete inaccurate or incomplete personal data that is part of the AI systems system or decision.

C. Require that data (evidence) analyzed, collected, or generated by AI systems comply with the rules of evidence.

See above.

D. Require using standards such as IEEE Std. 1012 for independent testing and auditing for any AI system procured by law enforcement and any AI system used in a criminal justice determination.

1. Independently verify and validate (IV&V) AI systems prior to deployment, or prior to informing criminal justice determinations in the legal system, law enforcement, governance, and related compliance.

The use of AI systems in criminal court can result in catastrophic failures through false imprisonment and the deprivation of people's rights. Scientists and engineers have long demanded that safety-critical software and hardware be the right systems built the right way. Therefore, AI systems should be independently verified and validated (IV&V) prior to deployment, or prior to informing criminal justice determinations in the legal system, law enforcement, governance, and related compliance. Specifically, AI systems ought to be independently verified and validated in accordance with technical standards such as IEEE 1012 Standard for System, Software, and Hardware Verification and Validation,[10] and be subject to recurring post-deployment audit, including with respect to their operators. We encourage the U.S. Government to uphold these same requirements.

IEEE1012 provides a universally applicable and broadly accepted process for helping to ensure that a product is correctly built for its intended use.[11] It is used to verify and validate Department of Defense nuclear weapons systems and NASA manned space systems and critical space exploration probes, among many others.

IV&V are interrelated and complementary processes that build quality into any system. Verification is focused on a product, providing objective evidence for whether the product conforms to requirements, standards, and practices. Validation is focused on customers and stakeholders, providing evidence for whether a product is accurate and effective, solves the right problem, and satisfies the intended use and user needs in the operational environment. In short, verification ensures that a product is correctly built, while validation ensures that the right product is built.

---

[10] IEEE Standard for System, Software, and Hardware Verification and Validation, IEEE Standard 1012-2016, Sept. 2017 (hereinafter referred to as IEEE 1012) (available at https://standards.ieee.org/ieee/1012/5609/). Other standards and requirements addressing software IV&V include IEEE Standard P3119 for the Procurement of Artificial Intelligence and Automated Decision Systems (available at https://standards.ieee.org/ieee/3119/10729/) and NASA Procedural Requirements 7150.2D "NASA Software Engineering Requirements" (available at https://nodis3.gsfc.nasa.gov/displayDir.cfm?t=NPR&c=7150&s=2D).

[11] See generally, Matthews, Jeanna. "How should we regulate AI: Practical Strategies for Regulation and Risk Management from the IEEE 1012 Standard for System, Software, and Hardware Verification and Validation." IEEE-USA (2023). (available at https://ieeeusa.org/product/how-should-we-regulate-ai/).

In the context of AI systems, IV&V answers the following types of questions: Is the analysis used by the AI systems software the best available, coded as designed, and appropriate for the problem? Does AI systems systematically favor including or excluding certain types of information? How likely are false negatives and false positives? Would outside experts agree with the software's results at each stage of analysis?

To help appropriately perform IV&V, IEEE 1012 requires that each software and hardware component be assigned an integrity level that increases depending on the likelihood and consequences of a failure: negligible, marginal, critical (causing "major and permanent injury, partial loss of mission, major system damage, or major financial or social loss," referred to as integrity level 3), and catastrophic (causing "loss of human life, complete mission failure, loss of system security and safety, or extensive financial or social loss," referred to as integrity level 4).[12] As the integrity level increases, so too does the intensity and rigor of the required IV&V tasks.

AI systems, like all software, should undergo IV&V according to its integrity level as defined by IEEE 1012. The likelihood of AI systems to cause wrongful convictions in the criminal legal system clearly constitutes catastrophic failure, and therefore should be held to the highest integrity level (level 4),[13] known as classical IV&V[14].

Classical IV&V must be "rigorously" independent to avoid conflicts of interest that could lead to catastrophic failure. To this end, IEEE 1012 requires technical, managerial, and financial IV&V when testing software and hardware where catastrophic consequences could occasionally occur and where critical consequences will probably occur.[15] Moreover, letting developers certify their own software is a clear conflict of interest, and the IEEE/Association for Computing Machinery Code of Ethics for Software Engineers is clear about the obligation of developers to manage competing aims.[16] Full definitions of technical, managerial, and financial independence from IEEE 1012 are below, but, in brief, the following must all be independent from the group that oversees the design and building of software: personnel, problem formulation, test and analysis tools for IV&V (technical), responsibility for IV&V (managerial), and control of the budget for IV&V (financial).[17]

[12] IEEE 1012, p. 196.
[13] IEEE 1012, p. 196.
[14] IEEE 1012, p. 199.
[15] IV&V with "rigorous" (the highest level) technical, management, and financial independence "is generally required for integrity level 4 (i.e., loss of life, loss of mission, significant social loss, or financial loss) through regulations and standards imposed on the system development" - where "IV&V responsibility is vested in an organization separate from the development organization." IEEE 1012, p. 199.
[16] D. Gotterbarn, K. Miller, and S. Rogerson, "Computer society and ACM approve software engineering code of ethics," Computer, vol. 32, no. 10, pp. 84–88, 1999. doi: 10.1109/MC.1999.796142.
[17] IEEE 1012, p. 198.

Specifically, technical independence "[r]equires the IV&V effort to use personnel who are not involved in the development of the system or its elements. The IV&V effort should formulate its own understanding of the problem and how the proposed system is solving the problem."[18] "Technical independence means that the IV&V effort uses or develops its own set of test and analysis tools separate from the developer's tools."[19] And if sharing tools is necessary, "IV&V conducts qualification tests on tools to assure that the common tools do not contain errors that may mask errors in the system being analyzed and tested."[20] This independence requires the exclusion of parties with a stake in the outcome. Forensic labs, whether part of the law enforcement agencies or not, have a shared interest in the outcomes and therefore, must be excluded from the analysis of the testing.

Managerial independence "[r]equires that the responsibility for the IV&V effort be vested in an organization separate from the development and program management organizations. Managerial independence also means that the IV&V effort independently selects the segments of the software, hardware, and system to analyze and test, chooses the IV&V techniques, defines the schedule of IV&V activities, and selects the specific technical issues and problems to act on."[21] The IV&V effort must be "allowed to submit to program management the IV&V results, anomalies, and findings without any restrictions (e.g., without requiring prior approval from the development group) or adverse pressures, direct or indirect, from the development group."[22] For this reason as well, forensic laboratories must be excluded from being responsible for the testing.

Financial independence "[r]equires that control of the IV&V budget be vested in an organization independent of the development organization. This independence prevents situations where the IV&V effort cannot complete its analysis or test or deliver timely results because funds have been diverted or adverse financial pressures or influences have been exerted."[23]

Two things are clear about classical IV&V as applied to AI systems in the criminal justice system: First, developmental or internal testing conducted by the owners or developers of the AI systems, law enforcement, or forensic laboratories are insufficient. They have fundamental conflicts of interest that violate the requirements of technical, managerial, and financial independence.

Second, peer-reviewed publications, while a priceless tool for scientific inquiry, are not a substitute, nor a valid approximation of IV&V when determining reliability or trustworthiness of a deployed system. Peer-reviewed publications form the foundation of scientific advancement, but peer reviewers of scientific publications are not tasked with answering questions like

[18] IEEE 1012, p. 198.
[19] IEEE 1012, p. 198.
[20] IEEE 1012, p. 198.
[21] IEEE 1012, p. 198.
[22] IEEE 1012, p. 198.
[23] IEEE 1012, p. 198.

"Should the AI systems software or results be admissible in court? Is the AI systems software fit for the evidence in this legal case?" Peer reviewers do not have access to the system itself and are not tasked with assessing its reliability. Peer reviewers are assessing whether a publication deserves the attention of the scientific community, whether the results described deserve the attention of other scientists. With respect to specific legal cases, any individual case could go well beyond the bounds of the published studies.

2. Test AI systems in accordance with standards that adhere to principles of due process, openness, consensus, balance, and right of appeal.

By definition, standards are "published documents that establish specifications and procedures designed to maximize the reliability of the materials, products, methods, and/or services people use every day."[24] They are the basis on which the safety and credibility of new products and markets are verified, making them fundamental to the modern economy.[25] Because standards have such a profound effect, standards-setting organizations (SSOs), like IEEE SA, have significant legal obligations regarding the standards they develop and the processes by which they craft those guidelines, including contract, intellectual property, and antitrust law.[26] Among the many U.S. Supreme Court opinions dealing with SSOs, there are two particularly relevant rules the organizations must abide by to avoid liability: fair processes and independence.[27]

As a result, the IEEE SA standards-development process follows a well-defined and documented path, from concept to completion, guided by a set of five basic principles and imperatives that ensure fairness and good practices during the development cycle.[28]

- Due process: having highly visible procedures for standards creation and following them.
- Openness: ensuring that all interested parties can participate and are not restricted to a particular type or category.
- Consensus: requiring a supermajority of a group to approve a draft standard (75% of the ballots must be returned, with 75% of them voting yes).
- Balance: ensuring that voting groups include all interested participants and avoid an overwhelming influence by any one party.
- Right of appeal: allowing anyone to appeal a standards development decision at any point, before or after approval.

[24] "Developing standards." IEEE Standards Association. https://standards.ieee.org/develop/index.html.
[25] "Developing standards." IEEE Standards Association. https://standards.ieee.org/develop/index.html.
[26] A. Updegrove. "Laws, cases and regulations in the essential guide to standards." ConsortiumInfo, 2013. https://www.consortiuminfo.org/essentialguide/laws.php.
[27] A. Updegrove. "Laws, cases and regulations in the essential guide to standards." ConsortiumInfo, 2013. https://www.consortiuminfo.org/essentialguide/laws.php.
[28] "Developing standards." IEEE Standards Association. https://standards.ieee.org/develop/index.html.

As a standard developed by IEEE SA, IEEE Std. 1012 complies with these principles. The U.S. Government should ensure that the procedures of standards development and adoption by various organizations intended to apply to AI systems in the criminal justice system, such as the NIST Organization of Scientific Area Committees (OSAC) for Forensic Science, also comply with these principles.

3. Establish standards and certifications for AI systems and their operators, and funding for recurring benchmarking exercises and independent testing and research.

Trustworthy systems must adhere to principles including effectiveness (system creators and operators shall provide evidence of the effectiveness and fitness for purpose), transparency (that the basis of a particular decision should always be discoverable), accountability (systems shall be created and operated to provide an unambiguous rationale for all decisions made), awareness of misuse (system creators shall guard against all potential misuses and risks in operation), and competence (system creators shall specify and operators shall adhere to the knowledge and skill required for safe and effective operation).[29]

Fully identifying and characterizing the limitations, failure modes and sequences, and error rates is critical to the trustworthy and reliable use of AI systems. The U.S. government should encourage and promote standards for reasonable testing and best practices for clear documentation of what testing has been done and what the results were. Critically, this testing must include the operators and individual forensic laboratories.

Governments should make the reports documenting the required IV&V and audits of their AI systems public. Furthermore, governments should encourage, develop, and update standards and certifications for AI systems and their operators, and fund recurring benchmarking exercises and independent studies to ensure their effectiveness, competence, inclusiveness, accountability, and transparency in operation. Specifically, we believe these standards, certifications, exercises, and studies should address:

- The requirements for informed trust by the general public in AI systems and the development of metrics that are immediately and easily accessible by experts and non-experts alike;
- The existence or absence of reliable and unbiased underlying scientific principles and methods in AI systems;
- The requirements for recurring testing and auditing of the operation of AI systems, including the operators, field conditions, testing data, environments, methodologies, and performance metrics;

[29] IEEE Ethically Aligned Design at 18.

- The requirements for publicly available documentation by developers and testers of AI systems, and of the use of AI systems in individual and aggregate cases and decisions;
- The requirements for certification or loss of certification of operators and AI systems, and for their validation for AI systems already in use;
- The requirements for individuals to be able to access, review, contest, and correct the data about them, to review and contest the decisions that affect them, and to request human review of such data and decisions;
- The requirements for operation in an ethical manner; and,
- The requirements for identifying and addressing vulnerabilities and threats to security, safety, and privacy such as spoofing, evasion attacks, transfer learning attacks, and data poisoning.

4. Test and evaluate AI systems for their trustworthiness in their operational environments and in legal proceedings.

IV&V is predicated on the value of testing technology in operational environments. No software or hardware is "generally" reliable -- any technology is only fit for certain purposes. Even technologies that are widely considered to be reliable have known failure modes. For example, cellular telephones are widely considered to be reliable but are not classified as "generally" reliable because they do not work effectively in tunnels or underground. Further, the desire for a technology to be classified as "generally" reliable rather than to consider its reliability in a particular case is misguided. A core premise of labeling a product or process as "well-engineered" is that these operating conditions are specifically defined, tested against predefined standards, and accompanied with estimated rates of failure. Systems like AI systems are engineered products incorporating scientific models and therefore require not only the perspective of researchers who have published proofs-of-concept but also engineers who have used product trials and operational testing and evaluation to demonstrate system performance in operating conditions, against predefined standards, and estimated rates of failure.

Therefore, any conclusion about the reliability and trustworthiness of AI systems cannot be effective if detached from an analysis of how the technology is used in legal proceedings - the AI system's ultimate operational environment. It is insufficient to base any conclusion merely on peer-reviewed and laboratory studies without comparing it to the criminal investigations and cases where AI systems are used to implicate people's life and liberty. Notwithstanding the concerns over peer-reviewed studies discussed above, if the types of data, environment, and users for the legal proceedings are not similar to those from the peer-reviewed or laboratory studies, the studies have little value.

Ultimately, the relevant scientific community for AI systems is the community of software engineers, computer scientists, and test and evaluation experts. AI systems used in the criminal

justice system cannot be considered generally accepted by the relevant scientific community without proper, classical IV&V by software engineers, computer scientists, and test and evaluation experts to the levels outlined in standards like IEEE Std. 1012.

As an example, probabilistic genotyping systems are AI systems used by forensic analysts to identify the likely people contributing to a mixture of DNA.[30] When assessing the reliability of probabilistic genotyping systems, courts and governments necessarily hear from experts in DNA. That is insufficient to properly address the reliability of the AI system.

This is analogous to assessing the safety and reliability of a car. A driver may be able to assess the performance as an end-user who operates the car and knows some of the some of the specific use-cases of what they want the car to do. However, to properly assess the safety and reliability of a car, it is necessary to have a mechanic's assessment as they have the expertise to actually inspect critical software and hardware systems (e.g., brakes, headlights, tires, steering, suspension, air bags, fuel systems, on-board diagnostics, and diagnostic trouble codes), diagnose and repair any issues -- far beyond a driver's knowledge, skill, experience, training, or education. A driver's statement that the car is safe and reliable to their experience is necessary but insufficient without the assessment from the mechanic based on their expertise.

5. Ensure accountability and transparency in procurement and contracting for AI systems.

To support awareness, access, and research on the existence, fairness, safety, security, privacy, and ethical and societal impacts of AI systems, there must be accountability and transparency in government procurement and contracting for AI systems. The Government should look to standards such as the IEEE P3119 draft standard for the Procurement of Artificial Intelligence and Automated Decision Systems which is designed to strengthen AI procurement approaches with due-diligence processes to ensure that agencies are critically evaluating the kinds of AI services and tools they acquire.[31] Specifically, the government should not procure AI systems that (i) require the governmental entity to indemnify vendors for any and all negative outcomes; (ii) do not adhere to the eight principles in IEEE's Ethically Aligned Design for creating and operating AI systems that further human values and ensure trustworthiness (as may be reflected in articulated guidelines, standards, certifications, audits, and other sound documentation);[32] (iii)

[30] See generally, J.M. Butler, H. Iyer, R. Press, M.K. Taylor, P.M. Vallone, S. Willis, DNA Mixture Interpretation: a NIST Scientific Foundation Review. NISTIR 8351-draft, 2021, https://doi.org/10.6028/NIST.IR.8351-draft.
[31] IEEE Standard P3119 for the Procurement of Artificial Intelligence and Automated Decision Systems (available at https://standards.ieee.org/ieee/3119/10729/). For further information, see, Gisele Waters and Cari Miller, "A How-To Guide on Acquiring AI Systems: The IEEE standard helps government agencies strengthen their purchasing requirements." IEEE Spectrum (Jan. 23, 2024) (available at https://spectrum.ieee.org/guide-on-acquiring-ai-systems); Gisele Waters and Cari Miller, "5 Ways to Strengthen the AI Acquisition Process: This IEEE standard covers how to evaluate vendors and negotiate contracts." IEEE Spectrum (Mar. 26, 2024) (available at https://spectrum.ieee.org/5-ways-strengthen-ai-acquisition).
[32] IEEE Ethically Aligned Design.

do not comply with federal, state, and local anti-discrimination laws; or, (iv) are shielded from independent validation and verification, and public review.

E. Fund the development of standards and certifications, testing and evaluation, certification, recurring benchmarking exercises and independent studies of AI systems, while promoting AI education and removing barriers to ensuring public awareness, access, and research on the existence of, and societal impacts of AI systems.

Technical assessments of reliability are not the sole determination of trustworthiness. There are eight principles for creating and operating systems that further human values and ensure trustworthiness: (i) human rights: systems shall be created and operated to respect, promote, and protect internationally recognized human rights; (ii) well-being: system creators shall adopt increased human well-being as a primary success criterion for development; (iii) data agency: system creators shall empower individuals with the ability to access and securely share their data, to maintain people's capacity to have control over their identity; (iv) effectiveness: system creators and operators shall provide evidence of the effectiveness and fitness for the purpose of systems; (v) transparency: the basis of a particular system decision should always be discoverable; (vi) accountability: systems shall be created and operated to provide an unambiguous rationale for all decisions made; (vii) awareness of misuse: system creators shall guard against all potential misuses and risks of systems in operation; and, (viii) competence: system creators shall specify and operators shall adhere to the knowledge and skill required for safe and effective operation.[33]

Therefore, we recommend testing to ensure trustworthiness. Below we list additional requirements for ensuring the trustworthiness of systems in general which includes the automated decision systems such as AI systems and many of the forensic technologies used today. If those providing or using AI systems or any other forensic technologies do not adhere to these requirements, then they should not be deemed trustworthy or fit for their use in determining or affecting people's rights and liberties.

To ensure trustworthiness of AI systems, and other forensic technologies, we believe that governments should:

1. Fund the testing, evaluation, certification, and investigation of AI systems, including the research and development of principles, methods, measures, and metrics.

The adoption and acceptance of AI systems requires developing and sustaining public confidence in their quality, reliability, and compliance with regulations and social norms. Increased government funding for government and independent third-party evaluation and certification of AI systems is essential to ensure efficacy, transparency, traceability, accountability, and

[33] IEEE Ethically Aligned Design.

competency. Development of design requirements, methods, metrics, and environments so that AI systems can be tested and evaluated for interactions with different systems is critical in the adoption and acceptance of AI systems. To this end, mechanisms must be developed for identifying and accounting for the features of AI systems that could cause current testing, evaluation, certification, and investigation methods to misinform decision makers or the public about the risk of system deployment or the causes of system malfunction.

2. Provide stakeholders (including parties to cases, academics, journalists, and other researchers) access to the AI systems for assessments of reliability, validity, and whether that information is fit-for-purpose.

Users are too often inappropriately denied access or forced to overcome improper and unnecessary barriers to access AI systems in order to determine the degree of reliability, validity, and whether that information is fit-for purpose. There are many more users of AI systems than merely forensic scientists or law enforcement. Independent testing of proprietary or government AI systems by litigants, academics, journalists, and other researchers is needed to ensure that AI systems are properly vetted and held accountable. Governments must clarify whether and how proprietary AI systems may be reverse engineered, modified, and evaluated under laws such as the Computer Fraud and Abuse Act and the anti-circumvention provision of the Digital Millennium Copyright Act, and rules of procedure and evidence. Moreover, governments must take steps to affirmatively promote awareness, access, research, and testing including:

- Ensuring accountability and transparency in government procurement and contracting for AI systems;
- Identifying and disclosing the AI systems used by the government;
- Adopting clear procedures relating to collection, usage, storage and sharing of personal information collected and used by AI systems;
- Providing constituents notice about AI systems decisions, explanations for those decisions, and processes for challenging decisions or data;
- Prohibiting contractual barriers where certain parties such as law enforcement or forensic laboratories are able to access AI systems for training or testing, while researchers, defendants, and other interested parties are denied similar access; and,
- Specifically, in legal disputes, tribunals should permit disclosure under appropriate protective orders of intellectual property related to AI systems when necessary to obtain evidence in compliance with other judicial requirements, including constitutional requirements, discovery laws, or subpoenas.

3. Remove barriers to parties' access to information needed to ascertain relevant evidence about and from AI systems in legal disputes.

Specifically, in legal disputes where judges, juries, and lawyers are the users of AI systems results, barriers to parties' access to information needed to ascertain relevant evidence about and from AI systems should be eliminated.[34] Intellectual property protections should not be used as a shield to prevent duly limited disclosure of information needed to ascertain whether AI systems meet acceptable standards of effectiveness, fairness, and safety. Specifically, in legal disputes, tribunals should permit disclosure under appropriate protective orders of intellectual property related to AI systems necessary to obtain evidence in compliance with other judicial requirements, including constitutional requirements, discovery laws, or subpoenas. Furthermore, laws, procedures, and public funding should not make it more difficult for non-government parties in legal disputes to develop, obtain expertise regarding, or gain access to evidence from AI systems than for government parties to do so.

[34] For example, when source code is ordered to be provided, "information needed" requires providing sufficient information for the recipient to build, run, and test the software themselves including, at minimum:

- All software dependencies including third-party code libraries, toolboxes, plugins, frameworks, and databases;
- Software engineering and development materials describing the development, deployment, and maintenance of the version(s) of the software system used in the instant case, including software engineering documents and build instructions;
- All records of software glitches, crashes, bugs, or errors encountered during the developmental validation study;
- Software version numbers of the components of the system used for the developmental validation study; and,
- All records of unexpected results, including false inclusions, false exclusions and the conditions under which the unexpected results were achieved.

When source code is ordered to be provided, "access" requires, at minimum, that the source code be made available for inspection, in a format allowing it to be reasonably reviewed, searched, and tested, during normal business hours or at other mutually agreeable and reasonable times, and at mutually agreeable and reasonable locations.