Request for Information (RFI) on Public and Private Sector Uses of Biometric Technologies: Responses

**DISCLAIMER:** Please note that the RFI public responses received and posted do not represent the views or opinions of the U.S. Government, the Office of Science and Technology Policy (OSTP), or any other Federal agencies or government entities. We bear no responsibility for the accuracy, legality, or content of these responses and the external links included in this document. Additionally, OSTP requested that submissions be limited to 10 pages or less. For submissions that exceeded that length, the posted responses include the components of the response that began before the 10-page limit.
AI and Criminal Procedure Rights

Brandon L. Garrett* and Cynthia Rudin**

Introduction

Today, as data-driven technologies have been implemented across a wide range of human activities, new warnings have been issued from a wide range of sources, academic, public policy, and government, regarding the dangers posed by artificial intelligence to society, democracy, and individual rights. The Federal Trade Commission (FTC) has described more detailed views concerning unfair and deceptive practices that rely on AI and impact consumers, and the FTC has taken action against a series of corporations regarding different types of algorithms.1 Several pieces of legislation that would regulate algorithms have been introduced in Congress, none of which has been enacted, but meanwhile, states have been active in considering and also adopting legislation regarding uses of AI. The White House Office of Science and Technology Policy (OSTP) has called for an “AI Bill of Rights.”2

Our statement responds to the OSTP call for submissions on that topic and we focus specifically on uses of AI in the criminal system.3 We write to reflect our own views as researchers, respectively, in law, scientific evidence, and constitutional law more broadly, and in artificial intelligence, machine learning, and computer science more broadly. We write to emphasize two basic points, that (1) artificial intelligence (AI) need not be black box and non-transparent in the ways in which it affects criminal procedure rights, and in fact, nothing will be lost by requiring such transparency through regulation; and (2) while more rights protections and regulations should be considered, far more can and should be done to apply and robustly protect the existing Bill of Rights in the U.S. Constitution as it should apply to uses by government of AI in the criminal system, particularly when AI is used to provide evidence regarding criminal defendants.

First, particularly in criminal cases in which life and liberty may be at stake, there should be a presumption that uses of AI directed towards providing evidence against criminal defendants, including by the federal government, such be fully interpretable and transparent. The burden to justify “black box” uses of AI in court should be a high one, given our commitment to public

---

* L. Neil Williams, Jr. Professor of Law, Duke University School of Law and Director, Wilson Center for Science and Justice.
* Professor of Computer Science, Electrical and Computer Engineering, Statistical Science, Mathematics, and Biostatistics & Bioinformatics, Duke University.
The views expressed here reflect only those of the authors and not those of any institution to which they belong, such as Duke University.
judicial proceedings and defense rights of access. There is no evidence that performance and efficiency depend on keeping the operation of AI secret from the public and unintelligible to users. That fundamental point, that AI can and should be open, for inspection, vetting, and explanation, is a simple one and it can be more forcefully insisted on at the federal level.

Second, we do not disagree that existing rights need to be at times reinterpreted for the AI era. However, we want to be sure that there is also a strong commitment to enforce existing constitutional criminal procedure rights, particularly given how difficult it is to amend the U.S. Constitution, but also given the unfortunate reality that those constitutional rights have been unevenly enforced in criminal cases, given the challenges that largely indigent defendants face in obtaining adequate discovery and the pressures to plead guilty and waive trial rights. The federal government in particular, should lead by example, its use of AI technologies, to vigorously protect constitutional rights of criminal defendants. In some settings, the federal government has already done so, but in others, the federal government has not taken individual rights concerns sufficiently seriously. We discuss below uses of AI that do not implicate constitutional criminal procedure rights to the same degree, and also highlight how crucial it is to focus on the uses to which AI is put during criminal investigations.

I. What is AI?

“Artificial intelligence” simply means that machines perform tasks that are typically performed by humans. Machine learning is a subfield of AI, and it heavily overlaps with predictive statistics. We should think of machine learning as a kind of pattern-mining, where algorithms are looking for patterns in data that can be useful. The data is supplied to the machine, which relies on past patterns to develop methods for making recommendations for what to do next. For instance, when predicting whether someone might have a drug overdose, patterns in their medical record and twitter feed, as well as those of others, might help us predict that outcome. These patterns can help human decision makers because no human can calculate patterns from large databases in their heads. Individual people may in fact be biased or place undue weight on information that is not particularly predictive. If we want humans to make better data-driven decisions, machine learning can help with that.

Simply put, machine learning methods can extract patterns from large databases that humans cannot. However, humans have a broader systems-level way of thinking about problems that is absent in AI.

The usefulness of AI as a tool in part depends on what data we feed to it. Just like a saw may perform irregularly if we feed it rotten wood, AI will perform poorly if we supply it with incomplete or irrelevant or biased data. If in the past, police often decided to arrest people simply based on their race, then relying on that policing data, AI will predict future arrests based on those same baked-in prejudices. If wealthier people have more access to certain medical services, then AI may recommend that medical support based on their past usage, and ignore others who may be in greater need of care.
II. Black Box Models Are Not More Accurate Than Interpretable "Glass Box" Models

First, what is black box AI? A black box predictive model is a formula that is too complicated for any human to understand or it is deemed by the designer to be proprietary, which means no one can understand its inner workings, because those inner workings are not shared or are not designed to be share-able. These models can cause problems for high stakes decisions like criminal risk scoring, where someone could get denied parole and they and their defense lawyer, the parole officers, and the public, are not able to figure out why the person did or did not get a high-risk score.

There is a common misconception that black box AI must be more accurate than any model a human could understand. That is just not true. Models that are interpretable to humans can perform just as well as models that are not. This has been shown to be true across fields, including computer vision. And recidivism risk scoring. The ways in which AI affects rights and interests need not be hidden or secret. AI need not be a black box to attain the accuracy of a black box.

In fact, Black Box AI tends to lead to less accurate decision-making, because such models are harder to troubleshoot and use in practice. Typographical errors in the input to black box recidivism prediction models has led to catastrophic errors in decision making, deeply affecting people's lives. This type of poor decision-making is a direct result of unnecessary secrecy, weighted in favor of companies that sell black box models to the justice system, rather than weighted towards those individuals in the justice system subjected to the decisions made from these models.

III. What Constitutional Criminal Procedure Rights are Implicated?

We now have far greater appreciation for the fact that AI can affect people’s lives in all sorts of important ways. These include applications in our criminal system. AI is already used in a host of criminal investigation, pretrial, and sentencing-related settings. For example, algorithms are used for risk scoring, in order to predict the risk that someone will commit a crime if they are released on bail, or given parole. Many states mandate that risk scores be used in various decisions, always to inform a judicial or other officials’ discretion, to be sure (and there are real questions concerning variability with which judges and others to incorporate quantitative information into their decision-making). Another high-profile example is the use of facial recognition technology as a forensic tool and for surveillance.

We emphasize throughout that the particular use of AI is important and can greatly alter the accuracy, privacy, and fairness interests at stake, as well as the fair trial rights involved. Thus, using AI to search for a missing person feared to have been kidnapped raises far fewer questions

---

5 Chaofan Chen, Oscar Li, Chaofan Tao, Alina Barnett, Jonathan Su, Cynthia Rudin, This Looks Like That: Deep Learning for Interpretable Image Recognition, NeurIPS, 2019.
than using AI to identify a culprit from a surveillance video. Any use of AI that results in evidence introduced during a criminal investigation, or in court, will generally raise far more constitutional concerns than a use of AI that is not used to prosecute a person.

A. Fair Trial Rights

A range of constitutional rights apply to protect individuals against deprivations of important interest through government action, and a range of rights are focused on the rights of individuals during criminal investigations and criminal adjudication. The most expansive constitutional provision implicated by uses of AI in criminal investigations strikes at the fundamentals of government action: the Due Process Clauses of the Fifth and Fourteenth Amendments. The federal government can ensure that no federal agency uses AI in a way that arbitrarily deprives persons of the rights during criminal investigations and adjudication. Simply put, we should look more closely at uses of AI that might result in evidence used to arrest a person and that might result in evidence used in court during a criminal case. We do not focus here on Fourth Amendment privacy rights relating to searches and seizures, although similar principles should apply. We focus here on due process and related rights: on uses of AI to generate evidence that could be used in court in a criminal case, to determine bail, to convict a person, or to impose a sentence.

The due process protections in criminal cases include assurances that all material and exculpatory evidence of innocence be disclosed to criminal defendants. Defendants have a right to effective assistance of counsel; defense counsel, in our view, cannot meaningfully defend a person without information about what AI evidence is being introduced in a case. The Equal Protection Clause protects against purposeful discrimination of protected groups, including based on race. The federal government can insist that AI be carefully vetted to assure against discriminatory impacts on minority groups. Further authority under civil rights legislation can assure that federal grant recipients do the same. Further, the defense cannot meaningfully defend a person without knowing whether the AI formula was calculated without error; in the case of risk scoring, there has been much evidence of typographical errors or other types of data errors influencing the scores. In some cases, it has been reported that the wrong score is being computed for all defendants: in the case of COMPAS in Broward County, FL, it was reported that the wrong scoring model had been used for years: the COMPAS parole score was used to determine pretrial risk, rather than the COMPAS pretrial score that was designed for this purpose.

---

12 How We Analyzed the COMPAS Recidivism Algorithm, Propublica, Jeff Larson, Surya Mattu, Lauren Kirchner and Julia Angwin, May 23, 2016
We emphasize the importance of affirmatively adopting policies to ensure that these constitutional rights are protected, however, because in practice, many are not meaningfully enforced. Discovery in criminal cases is typically quite limited, making it difficult for defendants to be aware that there is even an issue that exculpatory evidence may not have been disclosed. A criminal defendant may not be aware that AI was used to generate leads or evidence. Nor are evidentiary rights clearly defined in pretrial settings, or in sentencing proceedings in many jurisdictions. In general, expert evidence admissibility decisions have also been quite deferential in criminal cases; the National Academy of Sciences itself has explained that scientific safeguards must be put into place by government, given the limited ability of defendants to challenge even wholly unscientific expert evidence in criminal cases.\(^\text{14}\) That report highlighted how courts have routinely found admissible a range of forensic evidence of reliability that simply has not been established, where: “With the exception of nuclear DNA analysis, however, no forensic method has been rigorously shown to have the capacity to consistently, and with a high degree of certainty, demonstrate a connection between evidence and a specific individual or source.”\(^\text{15}\) Further, a criminal defendant, if indigent, may often be denied funds to retain an expert to examine AI technology used by a prosecution expert.\(^\text{16}\) The defendant may have no way to independently verify the work done, using AI, by government investigators.

Courts have tended to narrowly view defense requests for discovery regarding evidentiary uses of AI, as well as forensic evidence more broadly, in criminal cases. They have tended to more expansively view discovery requests only when errors have come to light and the judges have realized that there were important reasons why that evidence could have resulted in exculpatory information. Often those revelations occur years after a conviction and when it is too late to adequately provide relief.\(^\text{17}\) Further, typographical errors or other data errors that could occur in a defendant's record could easily influence a proprietary risk score calculation without detection, and, as we will discuss shortly, courts have upheld the rights of companies to protect such formulas.\(^\text{18}\)

**B. Examples from Criminal Law**

We know that humans can be biased, too punitive, too lenient, or inconsistent, and AI has the potential, if used consistent with principles of transparency, interpretability, and fairness, to improve on existing outcomes. In some settings, AI has the potential to better protect people’s rights. For example, judicial officers for decades have often followed cash-bail schedules (short cheat sheets, basically) quite robotically. If the person is arrested for a given charge, then bail is set at some cash level, say $2,000 or $10,000, if the judge mechanically follows the schedule. The person’s individual situation does not matter, apart from the arrest charges. The resulting jail


\(^{15}\) NAS Report, supra, at 7.

\(^{16}\) See Paul C. Giannelli & Sarah Antonucci, *Forensic Experts and Ineffective Assistance of Counsel*, 48 No. 6 CRIM L. BULLETIN 8 (2012).

\(^{17}\) NAS Report, supra, at 44-45 (describing audits and quality control failures at labs around the country).

Detentions are often wholly unnecessary and even counterproductive regarding public safety (pretrial detention can be crimogenic).\(^{19}\) We need to give judges better tools to make these decisions. So far, risk scoring has been used, although not always carefully considered by judges. AI has the potential at least, to introduce better approaches.

The black box problem in AI has become pressing in the area of risk assessment, however, as entire judicial systems have risk assessment schemes, but often without disclosing how they were created or what their basis was. While the types of information used in a risk tool may be made public, often the underlying methods, validation data, and studies are not. Most crucially, sometimes the assumptions behind how a person’s a level of risk gets categorized as “high” or “low” are not explained or justified. Concerns regarding transparency, interpretability, and fairness persist in those settings.

The most prominent legal challenge to a black box risk assessment program was brought in Wisconsin, where a defendant argued that it violated his due process and equal protection rights to base his sentence on an algorithm, marketed by a private company (called Northpointe), whose operation and validating information was not disclosed to him. In the State v. Loomis case, the Wisconsin Supreme Court dismissed these due process claims, emphasizing that judges have discretion when they consider the risk instrument.\(^{20}\) The Court did say that sentencing judges must be given written warnings about the risk tool, including cautioning judges that it relies on group data; those warnings do not open the black box in any way, however, or give judges any tools with which to judge the operation or accuracy of the AI for the individual person whose case is in front of them. Nor does it address the issue of possible typographical errors. And still, the defendant has no ability to view the formula or check its correctness or assess its applicability.

The federal government has put advance thought into ensuring more open uses of AI, when in the First Step Act, Congress legislated the use of risk assessments regarding federal prisoners. The Act called for a panel of researchers to vet the research design for this new risk assessment instrument, annual validation, and even “a requirement that [BOP staff] demonstrate competence in administering the System, including interrater reliability, on a biannual basis.”\(^{21}\) The legislative text was noteworthy in its embrace of a more open approach.

Unfortunately, after enactment, when First Step Act resulted in the development of the PATTERN risk assessment, the developers of the PATTERN, as well as the Department of Justice, in approving the risk instrument, did not explain the key choices: the selection of risk thresholds, or the validation data, which itself was not shared with other researchers. One problem was that the Act itself did not provide guidance on what should be deemed high, medium, low, or minimal risk. The Act provided even less information about how the dynamic or treatment related “needs” items should be operationalized, resulting in real concerns with the PATTERN instruments definitions of such items. The authors of the PATTERN have not shared annual data regarding


\(^{20}\) State v. Loomis, 881 N.W.2d 749, 767 (Wis. 2016).

the performance of the risk instrument, either. Only very general information has been reported, including that errors in the design were uncovered and supposedly corrected.\(^{22}\)

Second, a wide range of AI is now used in forensics, to conduct analyses on physical and biometric evidence from crime scenes. In forensics, we traditionally often had people who look at patterns and called a “match,” or a source identification, whether it was fingerprints, or firearms, or bitemarks. We know that they get it wrong and innocent people have been convicted based on those mistakes. AI may be able to improve on this pattern recognition work. Replacing humans with machines may not be bad if humans are comparatively more error prone. We need to be sure, though, that the machines work better and that they work fairly - or that they work at all.

To return to facial recognition technology, across the country, driver’s license photos are being fed into the federal face recognition system, along with other photos, such as images captured from airport cameras and the like.\(^{23}\) None of us agreed to have our faces included. We are part of an omnipresent lineup, and it is one maintained (in one such effort) by the federal government. The Federal Bureau of Investigation (FBI) maintains the FACE system of facial recognition. Its use raises privacy implications. It also raises accuracy questions. How likely is it that we will be misidentified? If a person is charged with a crime based on a “hit” using the federal FACE database, what can we say about how good the match is?

The FBI has been unwilling to share how the FACE algorithm works, what data it was trained on, and nor has the FBI detailed how the algorithm has been tested and how accurate it is. The GAO has repeatedly issued reports, given the FACE database use of large amounts of private biometric information, calling on the FBI to conduct such testing of false and negative positive rates.\(^{24}\) The FBI has responded that its policy “policy prohibits photos being provided as positive identification and photos cannot serve as the sole basis for law enforcement action,” and that ongoing work is being done to improve the accuracy of the system, including based on NIST.


\(^{23}\) Statement of Kimberly J. Del Greco, Criminal Justice Information Services Division Federal Bureau of Investigation Before The Committee on Oversight and Reform U.S. House of Representatives at a Hearing Concerning “The Use of Facial Recognition Technology by Government Entities and the Need For Oversight Of Government Use of This Technology Upon Civilians” 4 (2019) (“The FACE Services Unit performs facial recognition searches of FBI databases (e.g., FBI’s NGI-IPS), other federal databases (e.g., Department of State’s Visa Photo File, Department of Defense’s Automated Biometric Identification System, Department of State’s Passport Photo File), and State photo repositories (e.g., select State Departments of Motor Vehicles, criminal mugshots, corrections photos, etc.”), at https://docs.house.gov/meetings/GO/GO00/20190604/109578/HHRG-116-GO00-Wstate-DelGrecoK-20190604.pdf.

\(^{24}\) U.S. Government Accountability Office, Face Recognition Technology: DOJ and FBI Have Taken Some Actions in Response to GAO Recommendations to Ensure Privacy and Accuracy, But Additional Work Remains (2019), at https://www.gao.gov/products/gao-19-579t (“First, GAO found that the FBI conducted limited assessments of the accuracy of face recognition searches prior to accepting and deploying its face recognition system. The face recognition system automatically generates a list of photos containing the requested number of best matched photos. The FBI assessed accuracy when users requested a list of 50 possible matches, but did not test other list sizes. GAO recommended accuracy testing on different list sizes. Second, GAO found that FBI had not assessed the accuracy of face recognition systems operated by external partners, such as state or federal agencies, and recommended it take steps to determine whether external partner systems are sufficiently accurate for FBI’s use. The FBI has not taken action to address these recommendations.”).
evaluations. Hopefully federal and local law enforcement adhere to that restriction, and can improve the system, but it also begs the question whether such evidence should be used for preliminary criminal identification purposes, or as part (but not the sole) basis for a criminal prosecution, absent publicly-available information about its accuracy and operation.

If a facial recognition algorithm is used purely for an investigative purpose not designed to develop evidence against a suspect, such as to scan public places to search for a victim of human trafficking, then the same rights are not implicated. It is far more tolerable to use a tool that with less-clear evidence of reliability, purely as a way to generate leads to locate a missing person. The privacy rights of that missing person are not of salient concern. If a missing person is ultimately found based on those leads, then it is not relevant whether the system generated false leads or not, and nor do we typically need courtroom disclosure of how the system worked. However, the same system should not be used, without evidence of its reliability, to generate evidence linking a person to a crime. In the same way, police may rely on an anonymous tip of unknown reliability, to potentially generate leads in an investigation. If those tips help police locate a missing person or stolen property, then they reliability is corroborated, and there is little reason to inquire further into the source of the information. However, police cannot normally introduce statements by an anonymous tipster in court as evidence to support in a criminal prosecution.

There should be a strong presumption of transparency and interpretability for courtroom uses of AI. There may also be reasons to protect certain types of AI systems from disclosure, for which the presumption may be overcome. For example, if there is a strong national security justification for not making the full AI model public, at a minimum, it should be carefully vetted by independent researchers, with complete information about its strengths and limitations made available to users in law enforcement and the courts. This failure to open the black box on its FRT programs, shared with law enforcement around the country and used by federal agencies, is deeply troubling. We know quite a bit now about how accurate eyewitnesses are, and courts increasingly take real pains to ensure that a criminal jury hears about the reliability of eyewitness evidence, including due to concerns regarding error rates and cross-racial biases. Further, there is evidence that depending on what data they are trained on and how they are designed, that FRT can be racially biased unless (like Clearview AI's technology that uses biometric information from Internet users that have not given permission) it uses a massive dataset. NIST itself has reported on and documented real differences between facial recognition approaches, both regarding accuracy and bias. So far, courts have not ensured that the jury hears about how reliable untested FACE AI is. And that raises deep constitutional concerns as well as reliability concerns. We


26 Facial Recognition Technology: Current and Planned Uses by Federal Agencies (2021), at https://www.gao.gov/assets/gao-21-526.pdf (noting “18 of the 24 surveyed agencies reported using an FRT system, for one or more purposes”).

27 Dr. Charles Romine, Facial Recognition Technology (FRT), Testimony, Committee on Homeland Security, U.S. House of Representatives (2020) (noting “There, false positive differentials are much larger than for false negatives and exist across many, but not all, algorithms tested. Across demographics, false positives rates often vary by factors of 10 to beyond 100 times. False negatives tend to be more algorithm-specific, and often vary by factors below 3.”).
should not use AI or any other technique in order to identify suspects in criminal investigations if we do not know how good it is for achieving the purpose to which it is put.

This is an area where the federal government needs to lead in showing that use of AI robustly protects constitutional rights. Instead, the federal government is showing how readily it will permit defendant rights to be sacrificed in the name of expediency and profit by companies.

We note that our comments on surveillance to identify criminal suspects does not pertain to applications such as school security, where the goal is to eliminate a possible immediate threat. This is a separate topic than identifying suspects for criminal prosecution; they should not be confused or linked. For instance, it is possible to design security systems that require only biometric information from individuals who were previously identified as possible threats.  

Proposed Federal Legislation

It is noteworthy that the FTC has issued business guidance and begun enforcement regarding uses of AI in private industry, regarding non-transparent and misleading uses of AI and biased uses of AI, where they implicate consumer rights, under the FTC Act mandate to prevent unfair and deceptive practices. Each of those subjects should be also, as described, be the subject of federal efforts to prevent harms to the government’s own uses of AI in criminal cases. Similar efforts should be aimed at ensuring that government agencies do not violate constitutional criminal procedure rights, through non-transparent and unfair AI practices. We note that the U.S. House of Representatives has considered a “Justice in Forensic Algorithms Act” which would ensure that any algorithms used in criminal cases be unrestricted by any claim of proprietary or trade secrets protection, and vetted by NIST. Congressman Dwight Evans, D-PA, said: “Opening the secrets of these algorithms to people accused of crimes is just common sense and a matter of basic fairness and justice. People’s freedom from unjust imprisonment is at stake, and that’s far more important than any company’s claim of ‘trade secrets.’” Even absent such legislation, such an approach should be adopted by the federal government. Basic transparency standards and testing requirements should be follow by law enforcement and courts if they use AI tools in criminal cases.

Conclusion

We close by emphasizing that the burden of justification should be on those proposing to maintain non-transparent, black box use of AI. Government secrecy should never be the norm and the federal government should lead by example, and because due process should be understood to require it. The existing Bill of Rights provides important protections as against arbitrary action, without notice, and action that harms defendant’s fair trial and defense rights, as well as against discriminatory action in violation of the Equal Protection Clause and implementing civil rights acts. However, those protections have not proved effectual as remedies in criminal cases, given

---

28 Cynthia Rudin and Shawn Bushway, _A Truth Serum for your Personal Perspective on Facial Recognition Software in Law Enforcement_, Translational Criminology (2021).
limited pretrial discovery, inadequate defense resources, and a tradition of deferential gatekeeping regarding expert evidence. We ask that Office of Science and Technology Policy attend to these basic principles of open AI and careful and robust adherence to existing constitutional criminal procedure rights, as it conducts the important work of development of a broader AI Bill of Rights.