NLC CITY SUMMIT

NATIONAL LEAGUE OF CITIES SALT LAKE NOVEMBER 20–22, 2025

PRECONFERENCE SESSIONS AND EXECUTIVE EDUCATION NOVEMBER 19

Public Safety & Crime Prevention PSCP

Public Safety and Crime Prevention Committee

Agenda

(Subject to Change)

Wednesday, November 19, 2025
3:00 – 5:00 PM
Salt Palace Convention Center, Room 151 G – First Level

Chair: Councilmember Brian Kazy, Cleveland, OH
Vice Chair: Mayor Dontario Hardy, Kinston, NC
Vice Chair: Councilmember LaShay D. Harris, Rochester, NY

3:00 | Welcome

- Introductions, meeting overview, opening comments
- Pledge of Allegiance

3:05 – 3:15 PM | Federal Landscape

- Discussion of DOJ program impacts, FEMA Act of 2025, disaster reimbursement delays, preparedness grant allocations (UASI/SHSP), NFIP reauthorization, and resilience funding.

3:15 – 3:20 PM | Greetings from NLC Leadership

- **The Honorable Steve Patterson**, President, National League of Cities, Mayor, City of Athens, Ohio

3:20 – 3:35 PM | Federal Landscape (Continued)

3:35 – 4:05 PM | Presentation: The Prevent+Protect Program — Addressing the Mental Health Crisis in the First Responder Community

Presenters:

- Colleen Hilton, Founder of *AlliConnect* — Licensed therapist specializing in first responder mental health and former police spouse.

- Mayor Robert Matheny, Sr., City of Bridgeport, WV — Former Sheriff, Harrison County, WV; law enforcement professional for four decades; member of the Bearing Advisors team.

  Presenters will introduce the *Prevent+Protect Program*, which is designed to help cities strengthen the physical, mental, and financial health of their employees. The program's mental health component, delivered through AlliConnect, focuses on proactive mental fitness through peer support management, therapist matching, digital resources, wellness check-ins, and clinical support for government employers.

4:05 – 4:30 PM | Discussion: Artificial Intelligence in Law Enforcement and Real-Time Crime Centers

- Exploration of how AI and data analytics are transforming local public safety operations.
- Topics include ethical use, privacy, data governance, and workforce readiness for technology adoption.

4:30 – 4:40 PM | NLC Center for Municipal Practice's Justice Initiatives Work

Presenters:

- Tony McCright, Program Director, Justice Initiatives, NLC Center for Municipal Practice
- Maryam Ahmed, MPH, Program Manager, Justice Initiatives, NLC Center for Municipal Practice

NLC's Center for Municipal Practice provides research, education, and analysis on key topics and trends that impact the people in America's communities.

4:40 – 4:55 PM | 2026 PSCP Priorities and Workplan

- Discussion on committee priorities for 2026.

4:55 – 5:00 PM | Member Round-Robin and Next Steps

- Member updates on emerging public safety issues.
- Confirm follow-up items and adjourn.

Next PSCP In-Person Committee Meeting:
NLC Congressional City Conference
March 14-18, 2026
Washington, DC

NLC Procedures for the Adoption of National Municipal Policy and Resolutions

The National Municipal Policy (NMP) is NLC's comprehensive, standing statement of goals, principles, policies and program objectives on federal policy issues directly affecting or of concern to cities, towns and villages. The NMP serves as the basis for NLC's federal advocacy efforts on behalf of the nation's cities, towns and villages. The policy is subject to annual modification by delegates from direct member cities and state municipal leagues at the Annual Business Meeting during City Summit.

Since membership amends the NMP once each year, amendments to the policy typically do not endorse or oppose specific congressional bills, current presidential positions or technical aspects of federal regulations. Instead, positions on such timely matters – which are subject to major changes during the annual legislative and administrative processes – are the subject of NLC resolutions that stand for one year, from their time of passage until the adjournment of the next City Summit.

Direct member cities of NLC and state municipal leagues were invited to submit policy amendments and resolutions by June 27, 2025 for consideration by one of the seven Federal Advocacy Committees. The designated committee has the option of endorsing, amending or rejecting the submittal.

At least two weeks prior to the City Summit, proposed policy amendments and resolutions for 2026 are published on the NLC website and an announcement is sent to all NLC members. The proposed resolutions book for 2026 can be found here. These proposals are subject to change by the Resolutions Committee at City Summit prior to the Annual Business Meeting.

Federal Advocacy Committees

NLC's seven Federal Advocacy Committees manage NLC's National Municipal Policy and Resolutions. At the Congressional Cities Conference (CCC) in March, the Federal Advocacy Committees set agendas for the year. The Committees meet routinely, during the course of the year, to explore topics within their portfolio, engage in advocacy efforts, share best practices and develop policy and resolutions recommendations.

NLC's Federal Advocacy Committees met in person and virtually prior to City Summit to finalize their recommendations for policy amendments and resolutions. During these Federal Advocacy Committee meetings, any policy amendments or resolutions submitted to NLC by the June 27, 2025 submission deadline were considered. Committees had the option of endorsing or rejecting those submittals. Adoption of recommendations is by a majority vote of Federal Advocacy Committee members present and voting. Proxies are not permitted.

Resolutions Committee Meeting: Thursday, November 20

Proposals approved by the Federal Advocacy Committees are forwarded to the NLC Resolutions Committee for consideration. The Resolutions Committee will meet during City Summit on **Thursday, November 20 at 10:15 AM in the Hyatt Regency Salt Lake City, Salt**

Lake Ballroom CDE, 2nd Floor. The Federal Advocacy Committee chairs will report the recommendations of their respective committees to the Resolutions Committee members. The Resolutions Committee consists of the NLC Board of Directors plus representatives appointed by state municipal leagues whose states are not represented on the Board of Directors.

The Resolutions Committee will also consider any appeals from sponsors of proposals that were previously rejected by a Federal Advocacy Committee, as long as these were submitted by the June 27, 2025 deadline. Proposals introduced by members of the Resolutions Committee are also eligible for review.

Only members of the Resolutions Committee can participate and vote in this meeting. Speakers recognized during the meeting by the Chair may include Resolutions Committee members, Federal Advocacy Committee chairs or their designees, and sponsors of appealed policy recommendations. Decisions will be made by a majority vote of the members present. Proxies are not permitted.

The Resolution Committee Official Rules of Conduct and the NLC Bylaws shall govern the conduct of the Resolutions Committee meeting. In the event that procedural matters arise that are not addressed by the Official Rules or Bylaws, Robert's Rules of Order Newly Revised 12th Edition shall govern the conduct of the meeting.

Annual Business Meeting: Saturday, November 22

Resolutions Committee actions are referred to the Annual Business Meeting for consideration and adoption by the voting delegates. The report of the Resolutions Committee will include only recommended policy amendments and resolutions. The Annual Business Meeting will be held during City Summit on **Saturday, November 22 at 2:30 PM in the Salt Palace Convention Center, Grand Ballroom FHJ (Level One).**

To cast a vote at the Annual Business Meeting, all voting or alternate delegates must be present and registered with the Credentials Committee and must have official voting materials. Each direct member city has a certified voting delegate, or alternate, who is entitled to vote at the Annual Business Meeting. The delegate may cast a certain number of votes based upon the direct member city's population, determined by the 2020 U.S. Census; member cities may not split their votes. Each state municipal league is entitled to cast a total of 20 votes by its delegate or delegates, and those votes may be split and distributed at the discretion of each state municipal league. Voting delegates must be present to vote. Proxies are not permitted.

After a brief presentation of the Resolutions Committee's report, the Annual Business Meeting's Presiding Officer will call for adoption of NMP amendments and resolutions as proposed by the Resolutions Committee. Amendments to each chapter will be considered in the order in which those chapters appear in the NMP. Motions from the floor to amend the Resolutions Committee's recommendations require a majority vote for passage. Final adoption of amendments to the NMP requires a two-thirds vote of voting delegates.

Voting delegates may submit a petition for policy proposals to the NLC Federal Advocacy team by **10:00 AM on the day of the Annual Business Meeting** to the **Federal Advocacy Policy Office in the Salt Palace Convention Center, Room 255 B (Level Two).** Petitions must carry the text of the proposal and printed names, titles and signatures of 10 certified voting delegates

with their respective cities and states. The petition must receive a majority vote of the voting delegates to be accepted for floor consideration and require a two-thirds vote for final approval. Petitioners should complete the packet that can be found here.

The Official Rules of Conduct and the NLC Bylaws shall govern the conduct of the Annual Business Meeting. In the event that procedural matters arise that are not addressed by the Official Rules or Bylaws, Robert's Rules of Order Newly Revised 12th Edition shall govern the conduct of the meeting.

For further information about this process prior to City Summit or to contact the NLC staff for a Federal Advocacy Committee, contact Dion Taylor at 202-626-3064 or taylor@nlc.org.

During City Summit, please contact the Federal Advocacy staff at the Federal Advocacy Policy Office located in the Salt Palace Convention Center, Room 255 B (Level Two).

PSCP Committee 2025 Report Out

The Public Safety and Crime Prevention (PSCP) Committee advanced NLC's public safety and emergency management priorities throughout 2025 by engaging in key federal policy discussions and strengthening partnerships with Congress, federal agencies, and national organizations. The committee's work focused on ensuring that cities, towns, and villages are effectively supported in their efforts to build safe, resilient, and healthy communities.

Throughout the year, the committee contributed to NLC's advocacy on major legislative and regulatory actions affecting local governments. This included input on the FEMA Reform and Modernization Act of 2025 (H.R. 4669), which proposes significant updates to disaster cost share requirements, small disaster thresholds, and reimbursement procedures. PSCP members also monitored the implementation of Executive Order 14180, establishing the FEMA Review Council, and shared local perspectives to help ensure FEMA's modernization efforts reflect city needs. The committee supported reauthorization of the National Flood Insurance Program (NFIP) and continued to advocate for improvements in FEMA's Building Resilient Infrastructure and Communities (BRIC) and Hazard Mitigation Grant programs.

On public safety issues, the committee endorsed federal legislation to strengthen community based violence prevention, reentry, and mental health support for first responders. This included support for the Second Chance Reauthorization Act of 2025, as well as advocacy for continued funding for Department of Justice programs such as the Byrne Justice Assistance Grant (JAG), COPS Hiring, and Community Violence Intervention (CVI) initiatives. The committee also tracked key legal and policy developments related to policing and local law enforcement programs.

PSCP continued to emphasize the health and well being of first responders by supporting programs that promote peer support, mental health services, and proactive wellness initiatives. The committee also advanced NLC's advocacy for maintaining strong federal investments in fire and emergency response programs, including the Assistance to Firefighters (AFG) and SAFER grants, while highlighting ongoing local challenges related to equipment procurement and grant distribution.

During the Summer Board and Leadership Meeting, PSCP held a joint discussion with the Human Development (HD) Committee on the local impacts of immigration, focusing on the operational, fiscal, and human service challenges cities face as they support newcomers and coordinate with federal and state partners. This cross-committee collaboration provided valuable insight into how immigration intersects with local public safety, emergency management, and community resilience efforts.

Throughout the year, PSCP worked closely with NLC's other policy committees, state municipal leagues, and partner organizations including the U.S. Conference of Mayors, National Association of Counties, and the International Association of Fire Chiefs to ensure local perspectives were represented in national policy discussions.

Looking ahead, the PSCP Committee will continue to strengthen partnerships with FEMA, DOJ, and other federal agencies while identifying new opportunities to support local leaders on issues of disaster preparedness, public safety funding, and first responder wellness. The committee will also explore the role of artificial intelligence and emerging technologies in local law enforcement and public safety operations, with an emphasis on transparency, accountability, and community

trust. The committee's work this year reflects NLC's broader mission to reinforce a strong local federal partnership and advance policies that make America's communities safer, more resilient, and better prepared for future challenges.

NLC NATIONAL LEAGUE OF CITIES
A CENTURY OF STRENGTHENING CITIES

Briefing Sheet
Federal Policy Questions on AI Use in Local Law Enforcement

Purpose

To guide committee discussion on priority federal legislative and regulatory actions that would assist cities and law enforcement agencies in adopting AI technologies safely, equitably, and effectively.

Background

Local governments across the United States are increasingly considering or piloting the use of artificial intelligence (AI) in law enforcement operations, including surveillance, investigations, patrol management, and data analysis. While AI offers opportunities to improve efficiency and crime prevention, it also presents significant risks related to civil liberties, privacy, legal liability, and public trust.

The purpose of the discussion is to guide how NLC can support its member cities by advocating for federal policies that:

- Promote responsible AI adoption
- Minimize liability and risk for local agencies
- Ensure equity and transparency in public safety systems

NLC first adopted a resolution on local principles for the federal governance of generative artificial intelligence in 2023. The proposed resolution for NLC's 2026 National Municipal Policy follows this briefing sheet.

Discussion Questions for Federal Policy Development

1. Liability & Civil Rights Protection
    a. What uses of AI do you think pose the most potential risk for legal liability or civil rights violations? What protections or resources do cities need to avoid these risks?
    b. How can cities protect themselves from lawsuits related to misidentification, bias, or wrongful arrest caused by AI?
    c. To what extent should vendors of AI technology solutions be held liable for these lawsuits or civil rights violations, versus city governments?
2. Capacity Building

    a. What resources and supports do local police departments need to evaluate, procure, and manage AI technologies responsibly?
    b. Do local governments need technical assistance, vendor vetting standards, or regional AI advisors?
    c. What procurement policies could prevent cities from adopting untested or biased systems?
3. Privacy & Data Governance in Public Safety AI
    a. What privacy and data protection policies should be required of local law enforcement agencies using AI tools—especially those handling biometric, surveillance, or behavioral data? What should be required of AI vendors?
    b. Should there be a federal standard for how certain data may be retained by vendors or used for training AI models?
    c. What federal standards or testing requirements are necessary to ensure that AI-enabled surveillance complies with constitutional protections?
    d. Are you concerned about federal preemption of local decision-making authority over AI systems? If so, in what ways?
4. Access in AI Integration
    a. How can federal policy promote access to AI resources for smaller and mid-sized local governments?
    b. How can small and mid-sized cities be supported to adopt AI safely without widening public safety disparities?

Resources

1. NLC ITC Committee Resolution – Local Principles for the Governance of Generative Artificial Intelligence
2. NLC City Government Dashboard https://www.nlc.org/resource/city-ai-governance-dashboard/
3. NLC Artificial Intelligence Demystified- AI Toolkit for Municipalities: https://www.nlc.org/resource/ai-report-and-toolkit/
4. The Ethics and Governance of Generative AI https://www.nlc.org/article/2023/10/10/the-ethics-and-governance-of-generative-ai/
5. National Conference of State Legislatures – *"Artificial Intelligence and Law Enforcement: The Federal and State Landscape"* (NCSL) https://www.ncsl.org/civil-and-criminal-justice/artificial-intelligence-and-law-enforcement-the-federal-and-state-landscape
6. U.S. Department of Justice – *"Artificial Intelligence and Criminal Justice: Final Report"* https://www.justice.gov/olp/media/1381796/dl
7. Major Cities Chiefs Association (MCCA) – *AI Reference Guide (Feb 2024)* https://majorcitieschiefs.com/wp-content/uploads/2024/02/MCCA-AI_Reference-Guide-Feb-2024-.pdf
8. Police Chief Magazine (International Association of Police Chiefs)
    a. *"The AI-Policing Paradigm: Exploring the Benefits, Threats, Challenges, and Risks of Artificial Intelligence in Policing"* https://www.policechiefmagazine.org/ai-policing-paradigm/
    b. *"The Future of AI in Policing"* https://www.policechiefmagazine.org/the-future-of-ai-in-policing/
9. Interpol – *"Artificial Intelligence Toolkit"* https://www.interpol.int/en/How-we-work/Innovation/Artificial-Intelligence-Toolkit

NLC RESOLUTION 2025-47

LOCAL PRINCIPLES FOR THE GOVERNANCE OF GENERATIVE ARTIFICIAL INTELLIGENCE

ITC Committee Recommendation: Renew with Edits

WHEREAS, generative artificial intelligence (AI) technologies, which are algorithms capable of producing new outputs such as text, images, or other content, based on a set of training data, continue to advance at a rapid pace and are being deployed by federal, state, and local governments and private businesses for a variety of uses; and

WHEREAS, AI technologies have the potential to unlock new efficiencies and service models for local governments and spur local economic growth and activity; and

WHEREAS, AI technologies may also potentially negatively impact local economies and labor markets, government cybersecurity, equity, and resident vulnerability to criminal scams or disinformation; and

WHEREAS, Congress, the Administration, and nongovernmental organizations are currently exploring approaches to governing and regulating the public and private use of AI, including through America's AI Action Plan, ~~the Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence,[1]~~ the National Institute of Standards and Technology's AI Risk Management Framework,~~[2]~~ ~~the White House Office of Science and Technology Policy's Blueprint for an AI Bill of Rights,[3]~~ ~~the Senate SAFE Innovation Framework,[4]~~ the Bipartisan Senate AI Framework, and statements by the Institute for Electrical and Electronics Engineers;[5] and

NOW, THEREFORE, BE IT RESOLVED that Congress and the federal government should enact a regulatory framework that promotes innovation and investment in AI technology, while protecting local decision making, civil liberties, digital privacy, security, and transparency; and

BE IT FURTHER RESOLVED that technology solutions vendors should be required to provide full transparency to clients on their sources of training data for all AI-enabled tools, as well as the extent to which municipal client data is used, either jointly with other municipalities or independently, as further training data for AI tools; and

---

~~[1] The White House, "Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence," October 30, 2023.~~

~~[2] National Institute of Standards and Technology, "AI Risk Management Framework," January 26, 2023.~~

~~[3] The White House, "Blueprint for an AI Bill of Rights: Making Automated Systems Work for the American People," October 2022.~~

~~[4] Press release, "Majority Leader Schumer Delivers Remarks To Launch SAFE Innovation Framework For Artificial Intelligence At CSIS," June 21, 2023.~~

~~[5] IEE-USA, "Effective Governance of Artificial Intelligence," June 25, 2021.~~

BE IT FURTHER RESOLVED that federal standards should require vigorous detection and removal of bias within AI tools, as well as strict enforcement of these requirements to ensure human accountability for all decisions made in the development and use of AI; and

BE IT FURTHER RESOLVED that the federal government should require AI tools to be adequately tested for safety, reliability, performance and security according to valid third-party standards prior to being released for public use; and

BE IT FURTHER RESOLVED that a federal regulatory framework should protect public trust and privacy by requiring the disclosure of AI-generated or altered material, protecting the privacy of residents and the safety of children, requiring disclosure when user-generated or submitted content is used to train AI models, and preventing the promotion of mis- or disinformation; and

BE IT FURTHER RESOLVED that a federal regulatory framework must be developed with a focus on social responsibility and the input of all stakeholders, including communities likely to be most impacted by advancements in AI; and

BE IT FURTHER RESOLVED that local governments must retain the authority to decide on behalf of their own communities the extent to which they adopt artificial intelligence technologies for municipal operations, including piloting innovative new solutions and enacting local prohibitions on the usage of certain technologies by municipal agencies; and

BE IT FURTHER RESOLVE that local governments must not be preempted from managing the planning, permitting and siting of infrastructure necessary to support artificial intelligence, including but not limited to data centers, water and electrical services, and broadband infrastructure; and

BE IT FURTHER RESOLVED that Congress and the Administration should clearly define the legal rights, liabilities, and responsibilities involved in handling AI reference data, the ownership of reference data, and the requirements for disclosure, transparency, and accountability for AI models between the providers of AI tools and the users, including municipal governments; and

BE IT FURTHER RESOLVED that Congress should explore ways to foster regional economic growth, accelerate domestic research and development of AI and expand local government capacity for managing and deploying AI through incentives-based comprehensive funding models, such as that used in the CHIPS Act of 2022; and

BE IT FURTHER RESOLVED that Congress and federal agencies should proactively explore the impact of artificial intelligence on the employment market and invest in technical assistance and capacity-building, particularly for smaller local governments, assist negatively affected regions, upskill workers for changes in the job market, create an AI education pipeline, and encourage diversity within the AI workforce.

REPORT | ARTIFICIAL INTELLIGENCE, CYBERSECURITY AND PRIVACY TASK FORCE

Artificial Intelligence in Law Enforcement

The Federal and State Legislative Landscape

SAUL LOEB/AFP VIA GETTY IMAGES

NCSL
NATIONAL CONFERENCE OF STATE LEGISLATURES

JANUARY | 2025

Artificial Intelligence in Law Enforcement: The Federal and State Landscape

BY NICOLE EZEH, AMBER WIDGERY AND CHELSEA CANADA

The National Conference of State Legislatures is the bipartisan organization dedicated to serving the lawmakers and staffs of the nation's 50 states, its commonwealths and territories.

NCSL provides research, technical assistance and opportunities for policymakers to exchange ideas on the most pressing state issues, and is an effective and respected advocate for the interests of the states in the American federal system. Its objectives are:

- Improve the quality and effectiveness of state legislatures.
- Promote policy innovation and communication among state legislatures.
- Ensure state legislatures a strong, cohesive voice in the federal system.

The conference operates from offices in Denver, Colorado and Washington, D.C.

NATIONAL CONFERENCE OF STATE LEGISLATURES © 2025

JOAN CROS/NURPHOTO VIA GETTY IMAGES

Introduction

Technology and policing are closely intertwined; think speed detection radar technology and surveillance drones. Law enforcement officials use technology to better detect, investigate and solve crime. Simultaneously, concerns exist about efficacy and appropriate use. Artificial Intelligence is no exception.

Law enforcement agencies across the country are increasingly encountering and adopting technology equipped with AI. While officers investigate crimes that use AI, they also recognize that incorporating AI can increase efficiency and expand capabilities. AI governance is still in its infancy and law enforcement as well as state and federal policymakers are tasked with balancing the benefits of using AI with constitutional concerns.

As AI rapidly develops, policy aimed at promoting responsible use and education for law enforcement agencies will evolve. This issue brief provides examples of how federal, state and local law enforcement are incorporating AI into their work and reviews state and federal actions impacting the adoption.

AI Use by Law Enforcement

Law enforcement agencies use AI in three ways: to assist humans with tasks and increase capacity, expand human capabilities and, in some limited instances, replace humans entirely with fully automated processes.

While not a complete listing of technology or uses, the following exemplifies the myriad ways law enforcement and AI intersect.

MACHINE LEARNING

Machine learning trains algorithms to improve performance on tasks using data analysis. Law enforcement agencies generally possess a tremendous amount of data, such as information on arrests, location of crime, type of charges and clearance rates. Machine learning can be used to help predict where criminal activity is likely to happen and tailor decisions on staffing and type of response. This works by summarizing and communicating large data sets, enabling law enforcement agencies to make informed decisions on the use of limited resources and share trends with the community when appropriate.

COMPUTER VISION

Computer vision, a subset of machine learning, teaches computers to interpret and understand visual information from images or videos. Facial recognition is one of the most well-established uses, but it is also used for fingerprint matching, DNA analysis and ear biometrics.

Additionally, cameras enabled with computer vision can be used to look for specific objects and assist in enforcement of traffic laws. For example, security cameras can detect specific people, suspicious behavior and weapons. Traffic cameras can identify a stolen vehicle or enforce speed, red light and seat belt laws, acting as a force multiplier for law enforcement agencies.

Concerns with the use of this type of AI technology center around the issue of bias, such as difficulty detecting and distinguishing features of individuals with darker skin. There are AI redaction strategies that can reduce bias by removing characteristics of race and ethnic origin that may influence criminal charges.

NATURAL LANGUAGE PROCESSING

Natural language processing is another subset of machine learning that enables machines to understand, interpret and generate human language to, for example, enhance report writing. Reports need to be detailed and accurate, but they also take time. Some agencies have adopted technology that can write initial drafts of police reports based on body camera footage and real-time officer narration. Reports are then reviewed for accuracy and supplemented by officers.

Natural language processing is also used to review body camera footage to flag highly professional responses for purposes of recognition or identify problematic interactions for supervisor intervention.

AUTONOMOUS ROBOTICS

Autonomous robots are machines that can perform physical tasks in the real world. For example, drones are used for traffic collision reconstruction, arson investigations and investigations of other accidents such as train derailments and crane collapses. Some agencies are also experimenting with drones as a first response to calls for service, especially in situations where personnel safety would be at risk, such as armed suspects, car chases, hostage situations, bomb threats and missing persons searches on difficult terrain.

PROCESS AUTOMATION

Process automation is designed to assist with specific tasks and can be achieved using different types of AI. One of the ways this has been adopted is computer-aided dispatch which can help to triage calls and improve response times. Calls can also be analyzed in real time and can help with translation when callers speak other languages. Process automation uses machine learning algorithms by using neural networks and deep learning techniques to perform specific functions. It has also been used for court reminder systems that help to reduce failure to appear in court and subsequent warrants and arrests.

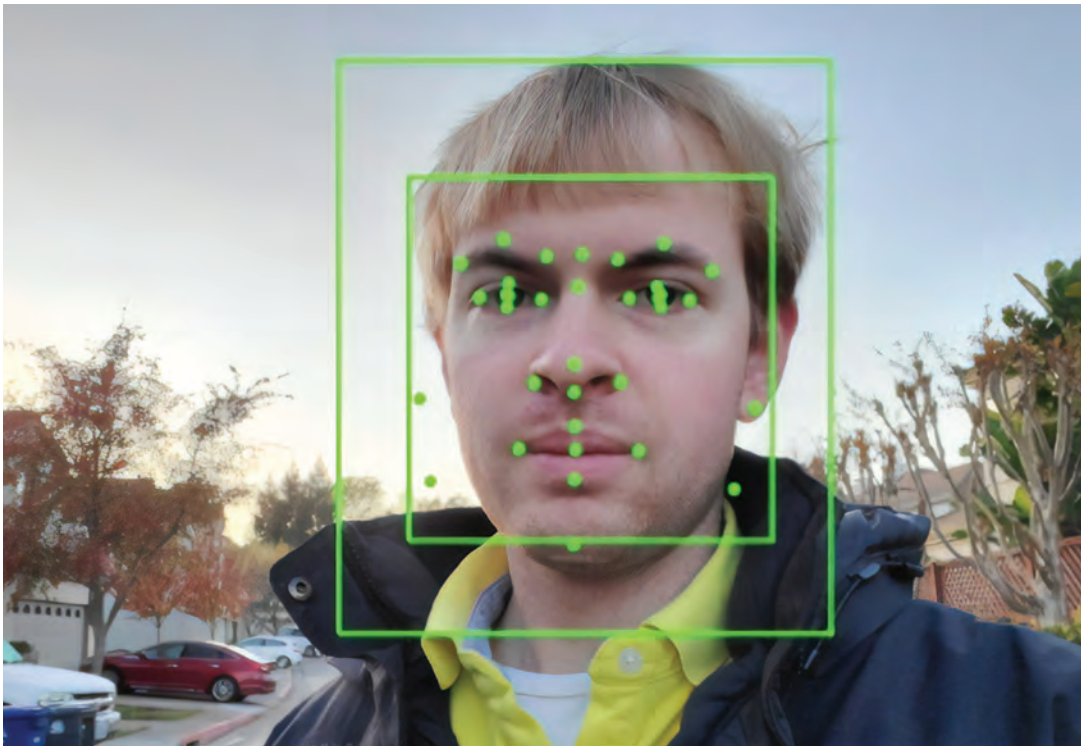
The Policy Landscape

Federal state and local governments can exercise oversight over policing, including adoption of AI technology. This is in addition to internal governance by law enforcement agencies themselves.

There are nearly 18,000 state and local law enforcement agencies across the country. The vast majority are categorized as local. Most police departments are led by a chief and report to a city mayor or a designated entity, like a police commission. Sheriffs, conversely, are generally elected at the county level and are accountable to voters with some budgetary oversight by county officials.

A state-level police agency exists in every state ranging in size from North Dakota's 139 officers to California's 7,202 officers. Legislatures provide annual funding for state police agencies such as highway patrols and bureaus of investigation.

At the federal level, there are 90 agencies that employ 136,815 full-time officers. While federal agencies sometimes differ greatly in scope and function from state and local agencies, the primary function of more than two-thirds of federal officers is criminal investigation.



SMITH COLLECTION/GADO/GETTY IMAGES

As law enforcement agencies move to adopt AI, lawmakers at all levels of government have considered policies that balance the benefits of AI in policing with potential risks. No state or locality has adopted a comprehensive set of laws governing law enforcement use of AI. In states that have legislated, the approach has been either to limit the adoption of AI or address how very specific applications of the technology can be used.

State Legislative Actions

State legislators in at least 30 states considered over 150 bills relating to government use of AI in 2024. Legislation addressed inventories to track the use of AI across state government, impact assessments, creating AI use guidelines, procurement standards and government oversight bodies. Some of these bills apply to government agencies broadly and may impact technologies used by state and local law enforcement agencies. As state legislatures continue to focus on AI regulation, states are also introducing legislation regulating the specific use of technology by law enforcement that have AI capabilities.

FACIAL RECOGNITION TECHNOLOGY

Over the last five years, at least 18 states have considered legislation to regulate law enforcement's use of facial recognition technology. In 2020, Washington was one of the first states to enact more comprehensive legislation regulating how state agencies and law enforcement use AI.

Washington's 2020 S 6280 and Colorado's 2022 S 113 require an accountability report, data management, security protocols, training procedures and testing, for government entities to use facial recognition technology. Additionally, entities must obtain a warrant or court order to use the technology to conduct ongoing surveillance, real-time identification or tracking. Utah enacted a law that prohibits government entities from using facial recognition on an image database except for law enforcement agencies. Agencies must submit a request and adhere to notice, data protection and disclosure requirements.

Some states have opted to temporarily or otherwise limit law enforcement use of FRT. In 2019, California enacted a three-year moratorium on use of facial recognition in body cameras. Oregon prohibits the use of facial recognition software captured by body cameras worn by law enforcement and New Hampshire limits use without proper authorization. Illinois enacted a law that prohibits law enforcement from using

DAVID MCNEW/GETTY IMAGES

drones equipped with facial recognition, while Vermont's 2021 legislation prohibits the use of facial recognition, except in cases involving sexual exploitation of children. Maine passed a law the same year prohibiting the search of "facial surveillance systems," with exceptions for serious crimes.

States have made clear that law enforcement cannot rely on facial recognition results as the single investigatory tool. Alabama now prohibits state and local agencies from using facial recognition as the sole basis for making an arrest or for establishing probable cause in a criminal investigation. Maryland authorized agencies to utilize facial recognition to establish probable cause or positive identification of an individual only if the results are supported by additional, independently obtainable evidence.

Other states convened study groups to provide recommendations for the use of facial recognition. Kentucky's 2022 legislation tasked a working group with creating a model policy for use by law enforcement agencies. That same year, Colorado created the Facial Recognition Task Force to investigate its use by state and local government agencies.

DRONE TECHNOLOGY

Drone capabilities are advancing as AI is incorporated into the technology. At least 15 states have passed laws that require law enforcement to obtain warrants before using drones. State legislatures continue to focus on the use of drones by law enforcement, and drone laws have been enacted in recent years.

In 2021, Florida expanded the authorized purposes for law enforcement drone use to include collecting evidence at crime scenes, assessment of damage during an emergency and vegetation or wildlife management. That same year Tennessee passed a law allowing officers to use drones for evidence collection; it was made permanent in 2023.

Illinois recently expanded its drone law to allow use by law enforcement at special events, to locate victims in an emergency and to conduct infrastructure inspection when requested by a local government. Utah now allows law enforcement to use drones in places that are off limits to others, like above certain critical infrastructure facilities.

States are addressing the potential cybersecurity risks associated with drones used by law enforcement. Tennessee passed a law that prohibits state agencies from purchasing equipment that meets criteria for

posing cybersecurity risks. In 2021, Florida passed a law that requires drones purchased by law enforcement to be from an approved manufacturers' list. Drones not from an approved manufacturer must be disconnected two years later. In 2024, Florida appropriated $25 million for drone replacement grants. Through this program, law enforcement agencies are required to provide the Florida Center for Cybersecurity within the University of South Florida the retired drones to analyze potential cybersecurity threats.

AUTOMATED LICENSE PLATE READERS.

At least 18 states enacted laws addressing the use of automated license plate readers or the retention of data collected by the automated readers. The devices capture computer-readable images that allow law enforcement to compare plate numbers against plates of stolen cars or cars driven by individuals suspected of criminal activity. They are mounted on police cars, road signs and traffic lights and capture thousands of plate images.

Data collected from the devices can enhance law enforcement's ability to investigate and enforce the law but also raises concerns about inaccurate information placed into databases and shared without restrictions on use, retained longer than necessary and used or abused in ways that could infringe on individuals' privacy.

Under Kansas law, a public agency is not required to disclose records that contain captured license plate data or the location of a license plate reader. Georgia law limits use of them to law enforcement purposes and required data to be destroyed no later than 30 months after it was collected, with limited exceptions. The law also mandates each agency maintain current policies regulating the use and train officers on appropriate use of the technology.

Federal Actions

CONGRESSIONAL ACTION

Congress and the Biden administration both contemplated the role of artificial intelligence technology in federal law enforcement crime response and investigation. Though the 118th Congress did not consider many bills on AI and policing, the Biden administration put out several policies and reports, including an executive order from the president. It remains to be seen whether the Trump administration will contin-

ue the efforts of the previous administration or pursue new approaches to artificial intelligence and law enforcement.

LEGISLATION INTRODUCED IN THE 118TH CONGRESS

- **HR 8005:** Child Exploitation and Artificial Intelligence Expert Commission Act of 2024. This bill establishes the Commission of Experts on Child Exploitation and AI to investigate and make recommendations to improve law enforcement's ability to detect, prevent, and prosecute AI-enabled child exploitation crimes. This was a bipartisan effort, with 15 Democrats and nine Republicans cosponsoring the bill.
- **HR 6143:** American Security Drone Act of 2023. This bill was passed through the National Defense Authorization Act for Fiscal Year 2024. It prohibits federal agencies from using drones manufactured in certain foreign countries to protect national security interests and foster U.S. manufacturing of drone technology.

OTHER CONGRESSIONAL ACTIONS AND REPORTS

The Government Accountability Office testified before Congress and released a report on federal law enforcement use of facial recognition technology. According to the report, seven federal agencies use the technology, including the Customs and Border Protection, Federal Bureau of Investigation, Secret Service, Bureau of Alcohol, Tobacco, Firearms, and Explosives, Drug Enforcement Agency, Homeland Security Investigations, and the U.S. Marshalls Service.

BIDEN ADMINISTRATIVE ACTIONS

The Biden administration released a plethora of agency reports beginning with President Joe Biden's Executive Order 14110: Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence in October 2023. The executive order directed federal agencies to produce reports and analyses on the use of AI by the federal government. In the law enforcement field, the order directed the Departments of Justice and Homeland Security to report the use of AI in the criminal justice system, consider how they can use their authorities to prevent algorithmic- and AI-related discrimination, and develop AI use recommendations for state and local law enforcement agencies.

The administration released several other reports and recommendations on law enforcement's use of artificial intelligence. In October 2024, the White House released a fact sheet outlining AI use restrictions and risk management practices in law enforcement necessary to safeguard national security. It also released a memorandum supporting the use of AI in a manner that fosters safety, security, and trustworthiness. Both documents emphasized the importance of protecting the constitutional rights of the public as well as those suspected of criminal activity. The memorandum highlighted the importance of not introducing biases based on protected characteristics and actions, such as race, ethnicity or participation in political speech.

Agency Reports and Guidelines

In December 2024, the Department of Justice released its final report in response to Executive Order 14110. The report discusses AI use in cases such as identification and surveillance, forensic analysis, predictive policing and risk assessment. The report includes best practices and measures to be taken before and after deploying AI technology. The report recommends that any criminal justice agency interested in using AI and creating an AI governance program should first identify the problem they wish to address and the reasons why the use of AI is preferable to non-AI alternatives. It highlights the importance of clear organizational structures for oversight, training and retaining a workforce with adequate resources to enact and enforce policies, and mitigating risks, among other recommendations.

In March 2024, the Department of Homeland Security released its Artificial Intelligence Roadmap, which outlines the agency's AI initiatives and the potential these initiatives have on homeland security as an enterprise. The document details the agency's approaches to AI use, such as ensuring all AI use cases meet due process requirements for legal proceedings, using AI to advance equity instead of amplifying existing societal inequities and taking a whole-of-government, collaborative approach with responsible use as a

guiding light. The report identifies opportunities for collaboration with and engagement from the private sector. The DHS maintains an AI use case library on its website.

TRUMP ADMINISTRATIVE ACTIONS

President Donald Trump issued many executive orders in the first week of his presidency. As of the publishing of this brief, two of these orders were related to artificial intelligence.

The first, EO 14148: Initial Rescissions of Harmful Executive Orders and Actions, rescinds Biden's EO 14110, which ordered the creation of the agencies reports outlined above. The second, Removing Barriers to American Leadership in Artificial Intelligence sets the Trump administration's AI policy moving forward. The order states "It is the policy of the United States to sustain and enhance America's global AI dominance in order to promote human flourishing, economic competitiveness, and national security." It also directs agency heads to review the reports, policies, and actions created pursuant to the revoked EO 14110 to make sure they are in line with the Trump administration's new policy. This leaves room for at least some of the work done under the Biden executive order to stay in place.

Conclusion

As AI technology continues to advance, government policies at all levels will likely evolve. Ongoing research into the effectiveness of AI applications in policing may have a part in shaping these policies. Research can also help develop better training programs to maximize the benefits of AI in law enforcement while minimizing the risks of misuse, bias and inaccuracies. Discussions about privacy, transparency and legal implications will likely remain central to an evolving landscape. Coordinated efforts across all levels of government may aid in the integration of AI-enabled technology in law enforcement that ensures responsible and effective use.

■ *Links to citations and legislation can be found in the online version of this publication.*
This publication and additional resources can be found at www.ncsl.org/civil-and-criminal-justice/artificial-intelligence-and-law-enforcement-the-federal-and-state-landscape

Nicole Ezeh is an associate legislative director in NCSL's State-Federal Relations Division.
Amber Widgery is a program principal for the Criminal and Civil Justice Program.
Chelsea Canada is a program principal in NCSL's Financial Services, Technology and Communications Program.

NCSL Contact:

dc-directors@ncsl.org

7700 East First Place, Denver, Colorado 80230, 303-364-7700 | 444 North Capitol Street, N.W., Suite 515, Washington, D.C. 20001, 202-624-5400

ncsl.org

© 2025 by the National Conference of State Legislatures. All rights reserved.

U.S. Department of JUSTICE

Artificial Intelligence and Criminal Justice

Final Report

December 3, 2024

Table of Contents

Foreword

Dear Mr. President,

The Department of Justice (DOJ), in consultation with the Department of Homeland Security (DHS) and the White House Office of Science and Technology Policy (OSTP), submits this report in response to Executive Order 14110, *Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence* (EO 14110).

Pursuant to the requirements set forth in EO 14110, this report satisfies the EO's directive in Section 7.1(b) that the Attorney General:

> submit to the President a report that addresses the use of AI in the criminal justice system, including any use in:
>
> (A) sentencing;
> (B) parole, supervised release, and probation;
> (C) bail, pretrial release, and pretrial detention;
> (D) risk assessments, including pretrial, earned time, and early release or transfer to home-confinement determinations;
> (E) police surveillance;
> (F) crime forecasting and predictive policing, including the ingestion of historical crime data into AI systems to predict high-density "hot spots";
> (G) prison-management tools; and
> (H) forensic analysis.

In each of those areas of the criminal justice system, this report "identif[ies] areas where AI can enhance law enforcement efficiency and accuracy, consistent with protections for privacy, civil rights, and civil liberties" and "recommend[s] best practices for law enforcement agencies, including safeguards and appropriate use limits for AI" and addresses the concerns set forth in both EO 14110 and Executive Order 14074, *Advancing Effective, Accountable Policing and Criminal Justice Practices to Enhance Public Trust and Public Safety* (EO 14074).

In addition, chapter IV of this report—Predictive Policing—is submitted in fulfillment of Section 13(e) of Executive Order 14074. This report satisfies EO 14074's directive that "the Attorney General, the Secretary of Homeland Security, and the Director of OSTP . . . jointly lead an interagency process regarding the use by [Law Enforcement Agencies] of facial recognition technology, other technologies using biometric information, and predictive algorithms."

Respectfully,

Merrick B. Garland
Attorney General

I. Introduction

Artificial intelligence (AI) use is rapidly transforming the criminal justice system and has the potential to make it more effective, equitable, and efficient. AI use also has the potential to cause harms, amplify disparities, and misdirect resources.

Today, AI in criminal justice predominantly involves conventional statistical analysis, such as regression models.[1] But that is changing. The accelerating pace of AI innovation is leading to increased use of computer vision, natural language processing, and other types of AI across the criminal justice system. Organizations and officials in the criminal justice system are also beginning to deploy generative AI systems.

The policy and technology choices that law enforcement agencies, pretrial and probation services, prison systems, and other criminal justice stakeholders make in the near term will affect millions of Americans. These choices will also set the trajectory for rapid expansion in the scope and scale of AI use throughout the criminal justice system.

On October 30, 2023, President Biden issued Executive Order 14110 on the *Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence* (EO 14110).[2] EO 14110 advances a coordinated governmentwide approach to responsible adoption of AI. It emphasizes that AI must advance equity and civil rights, respect privacy and civil liberties, and meet government performance objectives.

In order to "promote the equitable treatment of individuals and adhere to the Federal Government's fundamental obligation to ensure fair and impartial justice for all," section 7.1(b) of EO 14110 directs the Attorney General, in consultation with the Secretary of Homeland Security and Director of the Office of Science and Technology Policy (OSTP), to submit a report to the President on "the use of AI in the criminal justice system." The EO enumerates types of AI uses in criminal justice to address, and it further directs that the report "identify areas where AI can enhance law enforcement efficiency and accuracy, consistent with protections for privacy, civil

[1] This report uses a broad definition of the term "artificial intelligence," consistent with Section 238(g) of the John S. McCain National Defense Authorization Act for Fiscal Year 2019, Pub. L. No. 115-232, and the elaboration provided by OMB Memoranda M-24-10 and M-24-18 in implementing Executive Order 14110. The definition encompasses "[a]ny artificial system that performs tasks under varying and unpredictable circumstances without significant human oversight, or that can learn from experience and improve performance when exposed to data sets," as well as "[a]n artificial system developed in computer software, physical hardware, or other context that solves tasks requiring human-like perception, cognition, planning, learning, communication, or physical action," among other types of systems. Furthermore, "no system [is] too simple to qualify as covered AI due to a lack of technical complexity (e.g., the smaller number of parameters in a model, the type of model, or the amount of data used for training purposes)." MEM. FROM SHALANDA D. YOUNG, DIR., OFF. MGMT. & BUDGET, EXEC. OFF. OF THE PRESIDENT, TO HEADS OF EXEC. DEP'TS & AGENCIES (Mar. 28, 2024), available at https://www.whitehouse.gov/wp-content/uploads/2024/03/M-24-10-Advancing-Governance-Innovation-and-Risk-Management-for-Agency-Use-of-Artificial-Intelligence.pdf; MEM. FROM SHALANDA D. YOUNG, DIR., OFF. MGMT. & BUDGET, EXEC. OFF. OF THE PRESIDENT, TO HEADS OF EXEC. DEP'TS & AGENCIES (Sept 24, 2024), available at https://www.whitehouse.gov/wp-content/uploads/2024/10/M-24-18-AI-Acquisition-Memorandum.pdf.

[2] Exec. Order No. 14110, 88 Fed. Reg. 75191 (Oct. 2023), https://www.federalregister.gov/d/2023-24283.

rights, and civil liberties" and "recommend best practices for law enforcement agencies, including safeguards and appropriate use limits for AI."

The Department of Justice submitted an Executive Report in conformity with these requirements on October 29, 2024. This final report addresses these issues in richer detail and provides particularized recommendations.

Background

This report benefits from substantial input from stakeholders inside and outside of the federal government. Throughout 2024, Deputy Attorney General Lisa Monaco convened Justice AI, a series of six roundtable conversations that brought together law enforcement agencies and groups, civil society organizations, companies that develop AI products and services, and academic researchers who study AI use in criminal justice. In addition, the Civil Rights Division hosted four quarterly information exchanges with federal, state, and local agencies addressing civil rights issues associated with AI. The National Institute of Justice received comments for this report through a public request for input, and Department of Justice staff met with stakeholders throughout the process of preparing this report. The Department is grateful to the many experts who shared their valuable perspectives.

This report builds on a decade of U.S. Government initiatives that aim to ensure that AI use is effective, transparent, and respectful of privacy and civil liberties.

- In May 2014, the White House released a report entitled *Big Data: Seizing Opportunities, Preserving Values*.[3] The report touched on how algorithmic analysis of large datasets could have "tremendous" benefits for law enforcement activities, while also posing privacy and civil liberties issues.
- In May 2016, the White House issued the follow-up report, *Big Data: Algorithmic Systems, Opportunity, and Civil Rights*.[4] In a section on criminal justice, the report noted that uses of algorithms could advance public safety and public trust, but they must be "designed and deployed carefully" to prevent "exacerbat[ing] unwarranted disparities." The report also cautioned that "criminal justice data is notoriously poor" and often "inherently subjective."
- In October 2016, the National Science and Technology Council released the report *Preparing for the Future of Artificial Intelligence*, which summarized the state of AI in the government and made recommendations about AI governance and safety, noting

[3] EXEC. OFF. OF THE PRESIDENT, BIG DATA: SEIZING OPPORTUNITIES, PRESERVING VALUES (May 2014), https://obamawhitehouse.archives.gov/sites/default/files/docs/big_data_privacy_report_may_1_2014.pdf.
[4] EXEC. OFF. OF THE PRESIDENT, BIG DATA: A REPORT ON ALGORITHMIC SYSTEMS, OPPORTUNITY, AND CIVIL RIGHTS (May 2016), https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/2016_0504_data_discrimination.pdf.

that the lack of complete quality data in the criminal justice system risked "exacerbat[ing] problems of bias."[5]

- In December 2020, Executive Order 13960, *Promoting the Use of Trustworthy Artificial Intelligence in the Federal Government*, directed federal agencies to adhere to principles when using AI, including ensuring that AI applications are "consistent with the Constitution and all other applicable laws and policies, including those addressing privacy, civil rights, and civil liberties."[6] EO 13960 further directed that AI uses be "accurate, reliable, and effective" as well as "safe, secure, and resilient," "understandable," "regularly monitored," "transparent," and "accountable" through the implementation of "appropriate safeguards."

- Also in December 2020, Congress enacted the AI in Government Act, which required the Director of the Office of Management and Budget (OMB) to issue a memorandum that "identif[ies] best practices for identifying, assessing, and mitigating any discriminatory impact or bias on the basis of any classification protected under Federal nondiscrimination laws, or any unintended consequence of the use of artificial intelligence."[7]

- In January 2021, Congress enacted the National Artificial Intelligence Initiative Act of 2020, which established a National AI Advisory Committee under the Department of Commerce.[8]

- In May 2022, President Biden signed EO 14074, *Advancing Effective, Accountable Policing and Criminal Justice Practices to Enhance Public Trust and Public Safety*.[9] EO 14074 directed the Attorney General to commission a National Academy of Sciences (NAS) study on facial recognition, other biometric identification, and predictive algorithms used by law enforcement. NAS issued a report on facial recognition in January 2024 and convened a workshop on predictive policing in June 2024.[10]

---

[5] NAT'L SCI. & TECH. COUNCIL, COMM. ON TECH., EXEC. OFF. OF THE PRESIDENT, PREPARING FOR THE FUTURE OF ARTIFICIAL INTELLIGENCE 30 (Oct. 2016), https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf.

[6] Exec. Order No. 13960, 85 Fed. Reg. 78939, (Dec. 2020), https://www.federalregister.gov/documents/2020/12/08/2020-27065/promoting-the-use-of-trustworthy-artificial-intelligence-in-the-federal-government.

[7] Pub. L. No. 116-260, div. U, title 1, § 104 (codified at 40 U.S.C. § 11301 note), https://www.congress.gov/116/plaws/publ260/PLAW-116publ260.pdf.

[8] Pub. L. No. 116-617, div. C, title LI § 5104, https://www.congress.gov/116/crpt/hrpt617/CRPT-116hrpt617.pdf#page=1216.

[9] Exec. Order No. 14074, 87 Fed. Reg. 32945 (May. 2022), https://www.federalregister.gov/d/2022-11810.

[10] NAT'L ACAD. OF SCIS., ENG'G, & MED., FACIAL RECOGNITION TECHNOLOGY: CURRENT CAPABILITIES, FUTURE PROSPECTS, AND GOVERNANCE (Jan. 2024), https://nap.nationalacademies.org/catalog/27397/facial-recognition-technology-current-capabilities-future-prospects-and-governance; NAT'L ACAD. OF SCIS., ENG'G, & MED., LAW ENFORCEMENT USES OF PREDICTIVE POLICING APPROACHES (Nov. 2024), https://nap.nationalacademies.org/catalog/28037/law-enforcement-use-of-person-based-predictive-policing-approaches-proceedings..

- In October 2022, the White House Office of Science and Technology Policy (OSTP) published the *Blueprint for an AI Bill of Rights*, which recommended a set of principles for responsible AI use, including ensuring safety and efficacy, protecting against algorithmic discrimination, respecting privacy, providing notice and explanation, and establishing human oversight. The report noted that "[d]esigners, developers, and deployers of automated systems should take proactive and continuous measures to protect individuals and communities from algorithmic discrimination and to use and design systems in an equitable way."[11]

- In December 2022, Congress enacted the Advancing American AI Act, which directed the Secretary of Homeland Security to "issue policies and procedures … to ensure that full consideration is given to … the privacy, civil rights, and civil liberties impacts of artificial intelligence-enabled systems."[12] The Act also requires federal agencies to publish their AI use case inventories.

- In May 2023, the National Science and Technology Council released a revised *National Artificial Intelligence Research and Development Strategic Plan* to coordinate and focus federal investments in AI. The plan emphasized the importance of developing AI systems "in a manner that mitigates bias and harm and is done in accordance with the civil rights, civil liberties, and interests of those affected by the system."[13]

- In October 2023, President Biden signed EO 14110, which set out policy priorities and a whole-of-government approach for responsible AI development and implementation.[14] EO 14110 includes over 100 directives to agencies, including tasking the Attorney General with submitting this report.

- In March 2024, OMB issued Memorandum M-24-10, *Advancing Governance, Innovation, and Risk Management for Agency Use of Artificial Intelligence*. The memo fulfills a statutory requirement of the AI in Government Act of 2020 and a directive of EO 14110, and establishes baseline AI governance requirements for federal agencies, including governance structures, inventories, impact assessments, testing in real-world contexts, independent evaluation, consultation with impacted communities, and ongoing monitoring and risk mitigation.[15]

---

[11] EXEC. OFF. OF THE PRESIDENT, BLUEPRINT FOR AN AI BILL OF RIGHTS 5 (Oct. 2022), https://www.whitehouse.gov/wp-content/uploads/2022/10/Blueprint-for-an-AI-Bill-of-Rights.pdf.

[12] Pub. L. No. 117-263, div. G, title LXXII, subtitle B, §§ 7224(a), 7224(d)(1)(B), and 7225 (codified at 40 U.S.C. 11301 note), https://www.congress.gov/117/plaws/publ263/PLAW-117publ263.pdf.

[13] NAT'L SCI. & TECH. COUNCIL, SELECT COMM. ON A.I., EXEC. OFF. OF THE PRESIDENT, NATIONAL ARTIFICIAL INTELLIGENCE RESEARCH AND DEVELOPMENT STRATEGIC PLAN 2023 UPDATE 14 (May 2023), https://www.whitehouse.gov/wp-content/uploads/2023/05/National-Artificial-Intelligence-Research-and-Development-Strategic-Plan-2023-Update.pdf.

[14] Exec. Order No. 14110, 88 Fed. Reg. 75191.

[15] MEMORANDUM FROM SHALANDA D. YOUNG, DIR., OFF. MGMT. & BUDGET, EXEC. OFF. OF THE PRESIDENT, TO HEADS OF EXEC. DEP'TS & AGENCIES, (Mar. 28, 2024), available at https://www.whitehouse.gov/wp-content/uploads/2024/03/M-24-10-Advancing-Governance-Innovation-and-Risk-Management-for-Agency-Use-of-Artificial-Intelligence.pdf.

- In September 2024, OMB issued Memorandum M-24-18, *Advancing the Responsible Acquisition of Artificial Intelligence in Government*.[16] The memo complements M-24-10 with further acquisition guidance, including on how to enable testing of AI and ensure data sources are consistent with privacy and civil liberties protections, in fulfillment of section 7224(d) of the Advancing American AI Act.

This report also builds on work within the Department that has been central to the U.S. government's continuing efforts to ensure the responsible use of AI in criminal justice.

- In November 2023, Deputy Attorney General Monaco announced the establishment of DOJ's Emerging Technology Board (ETB), based on recommendations in the Deputy Attorney General's Comprehensive Cyber Review.[17]
- In February 2024, Attorney General Merrick Garland announced the designation of the Department's first Chief AI Officer (CAIO).[18] The CAIO and ETB are charged with developing and overseeing a comprehensive program of AI governance for DOJ, including implementation of EO 14110, the accompanying OMB Memoranda M-24-10 and M-24-18, and the National Security Memorandum on *Advancing the United States' Leadership in Artificial Intelligence; Harnessing Artificial Intelligence to Fulfill National Security Objectives; and Fostering the Safety, Security, and Trustworthiness of Artificial Intelligence*.
- In April 2024, the Chief of the Computer Crime and Intellectual Property Section within the Criminal Division sent a letter to the Copyright Office expressing support for a safe harbor for researchers who conduct independent bias testing on AI systems.[19] The letter noted that the Department benefits from this type of research in its work, including in the context of enforcement actions by the Civil Rights Division that are informed by this research.

---

[16] MEMORANDUM FROM SHALANDA D. YOUNG, DIR., OFF. MGMT. & BUDGET, EXEC. OFF. OF THE PRESIDENT, TO HEADS OF EXEC. DEP'TS & AGENCIES, (Sept. 24, 2024), available at https://www.whitehouse.gov/wp-content/uploads/2024/10/M-24-18-AI-Acquisition-Memorandum.pdf.

[17] Press Release, U.S. Dep't Just., Readout of Deputy Attorney General Lisa Monaco's Trip to New York and Connecticut (Nov. 9, 2023), https://www.justice.gov/opa/pr/readout-deputy-attorney-general-lisa-monacos-trip-new-york-and-connecticut.

[18] Press Release, U.S. Dep't Just., Attorney General Merrick B. Garland Designates Jonathan Mayer to Serve as the Justice Department's First Chief Science and Technology Advisor and Chief AI Officer (Feb. 22, 2024), https://www.justice.gov/opa/pr/attorney-general-merrick-b-garland-designates-jonathan-mayer-serve-justice-departments-first.

[19] Letter from John T. Lynch, Jr., Chief, Comput. Crime & Intell. Prop. Section, Crim. Div., Dep't Just., to Suzanne V. Wilson, Gen. Couns. & Assoc. Reg. Copyrights, Copyright Off., Libr. Cong, (Apr. 15, 2024), https://www.copyright.gov/1201/2024/USCO-letters/Letter%20from%20Department%20of%20Justice%20Criminal%20Division.pdf.

- Also in April 2024, DOJ joined five cabinet-level federal agencies in a pledge to uphold a national commitment to core principles of fairness, equality, and justice as use of AI and other emerging technologies continues to increase.[20]

- In September 2024, the Criminal Division released an update to its guidance on evaluating corporate compliance programs, setting an expectation that programs will address the role of AI in creating and identifying compliance risks.[21]

- In October 2024, the Criminal Division hosted a symposium on AI. The Principal Deputy Assistant Attorney General for the Division announced that the Department will be convening AI researchers and companies to understand how DOJ can best support independent AI accountability research.[22] This initiative follows steps the Department has previously taken to support security research, including a charging policy for good-faith research and guidance to industry.

- In October 2024, DOJ issued its plan for implementing a comprehensive AI governance program, consistent with OMB Memorandum M-24-10.[23] The plan describes DOJ's AI governance program, beginning with a thorough inventory of AI uses across the Department. For AI uses with heightened potential for impact on individuals' rights and safety, it provides for qualitative impact assessments, quantitative testing and ongoing monitoring for performance and biases, risk mitigation, and departmentwide coordinated decision-making.

The Use of AI in the Criminal Justice System

The types of AI uses in criminal justice described in EO 14110 fall into four categories, set forth below.[24] This report addresses each in turn, provides recommendations, and addresses the establishment of AI governance programs.

- **Identification and Surveillance.** From recognizing faces, fingerprints, and other biometric identifiers, to tracking license plates and locating gunshots, AI has a wide range of existing and potential applications for identification and surveillance in

[20] Press Release, U.S. Dep't Just., Five New Federal Agencies Join Justice Department in Pledge to Enforce Civil Rights Laws in Artificial Intelligence (Apr. 4, 2024), https://www.justice.gov/opa/pr/five-new-federal-agencies-join-justice-department-pledge-enforce-civil-rights-laws.
[21] U.S. Dep't Just., Crim. Div., Evaluation of Corporate Compliance Program (Sept. 2024), https://www.justice.gov/criminal/criminal-fraud/page/file/937501/dl.
[22] Press Release, U.S. Dep't Just., Readout of the Criminal Division's Symposium on Artificial Intelligence in the Justice Department (Oct. 3, 2024), https://www.justice.gov/opa/pr/readout-criminal-divisions-symposium-artificial-intelligence-justice-department.
[23] U.S. Dep't Just., Compliance Plan for OMB Memorandum M-24-10 (Oct. 2024), available at https://www.justice.gov/media/1373026/dl.
[24] Identification and Surveillance addresses "police surveillance" and "prison-management tools," as directed by EO 14110 sections 7.1(b)(i)(E) and (G). Forensic Analysis covers "forensic analysis," as required by section 7.1(b)(i)(H). Predictive Policing addresses "crime forecasting and predictive policing, including the ingestion of historical crime data into AI systems to predict high-density 'hot spots,'" as required by section 7.1(b)(i)(F). Risk Assessment covers "sentencing," "parole, supervised release, and probation, "bail, pretrial release, and pretrial detention," and "risk assessments, including pretrial, earned time, and early release or transfer to home-confinement determinations" as required by sections 7.1(b)(i)(A) through (D).

criminal justice contexts. These uses of AI can be significantly more accurate[25] and efficient than human observations and comparisons, and they can provide entirely new capabilities. But these AI uses also pose concerns, especially related to errors, bias, and privacy. When harms associated with these AI uses occur, they can be serious, including mistaken arrests, with a potential for disproportionate impact on certain communities. There is substantial nationwide variation in policies about whether and how AI may be used for identification and surveillance in criminal justice contexts.

- **Forensic Analysis.** AI can improve the capabilities, speed, and accuracy of forensic analysis. It is already being used to enhance DNA comparison, facilitate tracing of seized drugs, and prioritize electronic evidence, among other applications. Ongoing research suggests that future uses could include analysis of physical and trace evidence, medical evaluations, and assessing crime scenes. Forensic analysis must continue to meet exacting standards of accuracy and transparency to ensure due process and satisfy evidentiary requirements. Uses of AI may pose distinct challenges for meeting these requirements because of the complexity of validating and explaining AI-based forensic analysis, as well as the limitations of the data necessary for enabling these types of analyses.

- **Predictive Policing.** Law enforcement agencies use historical data to forecast the places where crime is likely to cluster and people who are at a higher risk of engaging in or being victims of criminal activity. Fundamental police work includes tracking where and when crimes occur, who is involved, and how crimes and people involved with crimes are connected. Developing accurate predictive models based on these types of data may help more efficiently direct resources—including non-law enforcement resources, such as social services—preventing crimes and decreasing response times. But there are also significant risks associated with predictive policing. The data used for predictive policing may have significant gaps and errors, and it may reflect human biases. Use of models based on that data may entrench existing disparities and result in unintended consequences and unjust outcomes, such as over-policing of certain individuals and communities. Successful place-based predictive policing programs integrate a range of strategies and interventions to promote public safety. At the same time, some law enforcement agencies have shifted away from person-based predictive policing, citing limited value and impact on privacy and civil liberties.

- **Risk Assessment.** Risk assessment tools estimate the likelihood that a certain individual outcome will occur in the criminal justice system, such as recidivating or failing to appear in court. These tools are widely used to inform pretrial release, sentencing, prison classification, probation, parole, and supervision. Used properly, risk assessment tools can be more accurate than human judgment alone and can enable more targeted use of various tools within the criminal justice system. Risk assessment tools can also be more transparent and equitable than human judgments. There are, however, significant risks associated with these tools. Risk assessment tools can be inaccurate, especially when they are not validated on local data or fail to take into

[25] This report uses the terms “accurate” and “reliable” as shorthand for the predictive performance of an AI system. The report uses more precise terminology when referring to specific performance metrics and measures, such as the precision or false positive rate of an AI system.

account relevant factors. They may be designed to estimate outcomes that are not directly relevant to the decision being made and thus fail to properly inform decision-makers. Risk assessment tools can perpetuate bias and inequality, since the data used in building models may reflect errors or biases and the development process may not incorporate input from affected communities. Models may also be unnecessarily complex, lack transparency, and apply substantially different categorizations to similar people.

II. Identification & Surveillance

Introduction

AI has the potential to enable agencies across the criminal justice system to more effectively and efficiently identify people based on biometrics, i.e., the measurement and analysis of individual physical characteristics.[1] Advances in computer vision, data mining, and complex pattern comparison tools, combined with decreasing costs of cameras and sensors as well as improvements in computation and data storage, have made AI-based biometric identification less expensive and more widely available.[2] Some agencies with law enforcement, correctional, and community supervision responsibilities now routinely use AI for biometrics.[3]

The use of AI for biometric identification also has risks. AI could misidentify individuals, which can misdirect law enforcement efforts and impact the civil rights and civil liberties of affected individuals. The performance of AI for biometric identification may also differ across demographic groups. There have been public reports of seven instances of mistaken arrests associated with the use of facial recognition technology, almost all involving Black individuals.[4] The collection and use of biometric data also poses privacy risks, especially when it involves personal information that people have shared in unrelated contexts.

The first part of the chapter opens with a description of biometric applications of AI in criminal justice, including automated fingerprint identification systems (AFIS), facial recognition technology (FRT), and iris scanning. The chapter then focuses on FRT use in criminal investigations to demonstrate the nuances of evaluating AI-based identification systems for accuracy and biases, as well as the importance of establishing policy frameworks and addressing impacts on privacy, civil rights, and civil liberties.

The second part of the chapter addresses uses of AI in the conduct of law enforcement surveillance. It focuses on automated license plate recognition (ALPR), an increasingly common practice for identifying vehicles that may be involved in criminal activity, including in ongoing emergencies.

---

[1] *See* INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, *Biometrics*, *in* INFORMATION TECHNOLOGY - VOCABULARY (3rd ed. Int'l Org. for Standardization 2022), https://www.iso.org/standard/73514.html ("automated recognition of individuals based on their biological and behavioural characteristics").

[2] Consistent with the definition of artificial intelligence in OMB Memoranda M-24-10 and M-24-18, this chapter categorizes identification and surveillance methods that involve algorithms or statistical analysis as AI.

[3] The following discussion of different biometric approaches includes references to examples of how these technologies are used by federal, state, and local agencies.

[4] U.S. COMM'N CIV. RTS., THE CIVIL RIGHTS IMPLICATIONS OF THE FEDERAL USE OF FACIAL RECOGNITION TECHNOLOGY 25 (2024), https://www.usccr.gov/files/2024-09/civil-rights-implications-of-frt_0.pdf; *see also* NAT'L ACADS. SCIS., ENG'G, & MED., *Facial Recognition Technology: Current Capabilities, Future Prospects, and Governance* 83 (2024), https://doi.org/10.17226/27397 ("NAS 2024 Report") (describing six then-known cases of mistaken arrests).

Biometrics

a. Automated Fingerprint Identification Systems

Law enforcement has used forms of fingerprint analysis for over a century.[5] The two common types of analysis today involve: (1) fingerprints collected from a known source in a controlled environment, commonly referred to as ten-prints, and (2) fingerprints from an unknown source collected from a surface or object, commonly referred to as latent fingerprints.[6] Analysis generally involves capturing friction ridge patterns on a person's skin, observing features including the locations and types of ridges (minutiae), and then comparing features across fingerprints for similarity.[7]

Automated methods for fingerprint comparison became practical in the 1970s and entered widespread use by the 1990s.[8] Live scan systems, which capture digital images of fingerprints without applying ink, developed in parallel. While identification based on fingerprints generally involves specialized scanning hardware or high-resolution digital cameras, technology may eventually enable ordinary smartphone cameras to capture latent prints.[9] Machine learning methods may also enable more advanced forms of fingerprint analysis in the future, such as comparing fingerprints for whether they may be from different fingers of the same person.[10]

Comparison of ten-prints is highly automated in practice today, often not involving a human examiner unless fingerprints will be offered as forensic evidence in a prosecution. Analysis

---

[5] *See* Press Release, Fed. Bureau Investigation, Crim. Just. Info. Serv. Div., FBI's Criminal Justice Information Services Division Celebrates 100th Anniversary of National Fingerprint Repository (July 10, 2024), https://www.fbi.gov/news/press-releases/fbi-s-criminal-justice-information-services-division-celebrates-100th-anniversary-of-national-fingerprint-repository ("In 1924, the FBI established an Identification Division informally called 'ident' for many years. 'Ident' gathered prints from police agencies nationwide and manually searched them upon request for matches to criminals and crime evidence.").

[6] Fingerprints may include impressions of palms, in addition to individual fingers.

[7] *See* John R. Vanderkolk, *Chapter 9: Examination Process*, *in* NATIONAL INSTITUTE OF JUSTICE, THE FINGERPRINTING SOURCEBOOK 9-13 (Nat'l Inst. of Justice 2011), https://www.ojp.gov/pdffiles1/nij/225329.pdf ("The direct or side-by-side comparison of friction ridge details to determine whether the details in two prints are in agreement based upon similarity, sequence, and spatial relationship occurs in the comparison phase.").

[8] *See generally* Kenneth R. Moses, *Chapter 6: Automated Fingerprint Identification System (AFIS)*, *in* NATIONAL INSTITUTE OF JUSTICE, THE FINGERPRINTING SOURCEBOOK 6-1 (Nat'l Inst. Of Justice 2011), https://www.ojp.gov/pdffiles1/nij/225326.pdf.

[9] Robert Pitts et al., *Empirical Comparison of DSLRs and Smartphone Cameras for Latent Prints Photography*, 3 WIREs FORENSIC SCI. (2021), https://wires.onlinelibrary.wiley.com/doi/epdf/10.1002/wfs2.1391 ("argu[ing] that the cameras equipped in current and future mobile devices are adequate for the purpose of latent print documentation and identification, making it a useful complement, if not a replacement, to DSLRs currently used by crime scene investigators and fingerprint examiners."); Maryah E. M. Haertel, Eduardo J. Linhares & Andre L. de Melo, *Smartphones for Latent Fingerprint Processing and Photography: A Revolution in Forensic Science*, 3 WIREs FORENSIC SCI. 2 (2021), https://wires.onlinelibrary.wiley.com/doi/epdf/10.1002/wfs2.1410 ("The use of smartphones in the search and acquisition of latent fingerprints is still new, but various studies show its possibilities.").

[10] Gabe Guo et al., *Unveiling Intra-person Fingerprint Similarity via Deep Contrastive Learning*, 10 SCI. ADVANCES (2024), https://www.science.org/doi/10.1126/sciadv.adi0329 (stating that "fingerprints from different fingers of the same person share very strong similarities" and suggesting that intra-person fingerprint similarities "can also help narrow down the candidate list generated by automated fingerprint identification systems.").

of latent prints, by contrast, usually involves review by a trained examiner because prints may be incomplete or degraded.[11] Automated methods may generate possible leads for subsequent analysis and support examiners in comparing fingerprints.

Fingerprint comparison is widely used for other purposes in the criminal justice system. Fingerprint-based checks are the standard for background checks in criminal justice, as well as for other positions of public trust, including teachers, childcare workers, and those in other sensitive occupations. Fingerprints are also the standard for identification based on criminal history record information in the United States. Fingerprint-based verification using 1-2 fingers may be used in certain applications to confirm the identity of an authorized user, track chain-of-custody of certain types of evidence, or limit access to sensitive areas.[12]

The FBI's Next Generation Identification (NGI) system provides fingerprint services to law enforcement agencies nationwide.[13] NGI contains over 217 million unique fingerprint identity records, over 28 million unique palm print identity records, and over 1.2 million unidentified latent prints. Almost every (if not every) state has its own automated fingerprint identification system, and these systems are also common at local law enforcement agencies.[14]

b. Facial Recognition Technology

Facial recognition technology uses methods from computer vision and other areas of AI to isolate and compare faces in photos or video. FRT became available for criminal justice use in the 2000s and became more widely used in the 2010s.[15] Algorithms have rapidly advanced in recent

---

[11] There is ongoing research and debate about statistical models to estimate the likelihood of fingerprint features from population base rates and validation of fingerprint comparison as practiced in particular laboratories. *See* PRESIDENT'S COUNCIL OF ADVISORS ON SCI. & TECH., FORENSIC SCIENCE IN CRIMINAL COURTS: ENSURING SCIENTIFIC VALIDITY OF FEATURE-COMPARISON METHODS (Sept. 2016), https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_forensic_science_report_final.pdf; William Thompson et al., *Latent Fingerprint Examination, in* AAAS, FORENSIC SCIENCE ASSESSMENTS: A QUALITY AND GAP ANALYSIS (AAAS 2017), https://www.aaas.org/sites/default/files/reports/Latent%20Fingerprint%20Report%20FINAL%209_14.pdf; Bradford T. Ulery et al., *Accuracy and Reliability of Forensic Latent Fingerprint Decisions*, 108 PROCS. NAT'L ACAD. SCIS. 7733 (2011), https://www.pnas.org/doi/full/10.1073/pnas.1018707108.

[12] The fingerprint readers used in these applications often lack the resolution necessary for criminal investigative uses of fingerprint analysis. As discussed below, there is a significant distinction between the 1:N matching that is common in investigations and the 1:1 matching that is common for these applications.

[13] FBI, FY 2025 President's Budget Request 55 (Mar. 2024), https://www.justice.gov/d9/2024-03/fbi_fy_2025_presidents_budget_narrative_3-5-24_final_1.pdf ("The NGI System services connectivity for 106,981 Federal, State, local, and Tribal law enforcement customers. These customers have existing statutory authorization to conduct background checks using the NGI System; however, only about one third (38,108) of those regularly do.").

[14] Nat'l Inst. of Justice, Latent Fingerprint Interoperability Survey: A National Study of Automated Fingerprint Information Systems (AFIS) Maintained by Law Enforcement Agencies 36 (2014), https://www.ojp.gov/pdffiles1/nij/247910.pdf. (showing a map of AFIS vendor information for state agencies which excluded Vermont (did not provide an answer), Minnesota and the District of Columbia (did not participate)).

[15] *See generally* Nat'l Inst. of Justice, History of NIJ Support for Face Recognition Technology (Mar. 5, 2020) at https://nij.ojp.gov/topics/articles/history-nij-support-face-recognition-technology (discussing NIJ role in face algorithm research and development); Statement of Jerome M. Pender Before the Senate Judiciary Committee, Subcommittee on Privacy, Technology, and the Law 112 Cong. (2012), https://archives.fbi.gov/archives/news/testimony/what-facial-recognition-technology-means-for-privacy-and-civil-

years, significantly increasing the ability of FRT systems used in criminal justice to correctly match faces.[16]

In the usual design of an FRT system, an algorithm first detects a person's face in a photo or video and extracts the relevant region.[17] Next, it computes a quantitative representation of the face (the "template"). Finally, the algorithm compares templates, producing a "similarity score" for pairs of templates. Some recent FRT systems use deep learning models that integrate these steps.[18]

There are, broadly, two types of FRT uses. One-to-one ("1:1") FRT compares a captured ("probe") image to a single other image or template, typically to verify a person's identity. One-to-many (also called "one-to-n" or "1:N") FRT compares a captured image to a database ("gallery") of known images.[19]

In the criminal justice system, one-to-one FRT has several applications. The Federal Bureau of Prisons uses one-to-one FRT to confirm employees' identities before entering secure areas of a facility.[20] Probation services may use FRT to allow individuals under court-ordered supervision to verify their identity via smartphone rather than requiring physical contact with a probation or pretrial officer.[21] Similarly, Customs and Border Protection and the Transportation Security Administration use one-to-one FRT to confirm traveler identities.[22]

---

liberties (discussing advances in the NGI program through July 2012); U.S. Gov. Facial Recognition Legal Series (Aug. 31, 2011), https://ucr.fbi.gov/fingerprints_biometrics/biometric-center-of-excellence/files/Forum_1_Minutes.pdf; William Casey et al., *Facial Recognition Technology — Baselining Uses and Legal Challenges: Meeting Minutes*, *in* U.S. GOVERNMENT FACIAL RECOGNITION LEGAL SERIES (2011), https://ucr.fbi.gov/fingerprints_biometrics/biometric-center-of-excellence/files/Forum_1_Minutes.pdf (noting advances made in facial recognition technology through August 2011). The FBI, for example, began developing the NGI-IPS system in 2008 and began using and providing access to the system in 2011. NGI-IPS became fully operational in 2015, at which point 7 states had access to the system. *See* GOV'T ACCOUNTABILITY OFF., PUB. NO. GAO-19-579T FACE RECOGNITION TECHNOLOGY: DOJ AND FBI HAVE TAKEN SOME ACTIONS IN RESPONSE TO GAO RECOMMENDATIONS TO ENSURE PRIVACY AND ACCURACY, BUT ADDITIONAL WORK REMAINS (June 2019), https://www.gao.gov/assets/gao-19-579t.pdf.

[16] The National Institute of Standards and Technology has performed FRT algorithm evaluations for over 30 years, and currently publishes FRT algorithm evaluations on an ongoing basis through the Facial Recognition Technical Evaluation (FRTE) program. In 2024, FRT algorithms commonly have a below 1% false negative rate with a false positive rate of 3 in 1,000. By comparison, the best performing algorithm in NIST's 2017 challenge had a 22% false negative rate with a false positive rate of 1 in 1000. *Face Technology Evaluation – FTRE/FATE*, Nat'l Inst. Standards & Tech, https://www.nist.gov/programs-projects/face-technology-evaluations-frtefate.

[17] *See* NAS 2024 Report *supra* note 4, at 32-34.

[18] *See* Mei Wang & Weihong Deng, *Deep Face Recognition: A Survey*, 429 NEUROCOMPUTING 215 (2021), https://doi.org/10.1016/j.neucom.2020.10.081 (discussing the emergence of deep learning models in FRT in 2012).

[19] One-to-one and one-to-many FRT systems are closely related, because one-to-many FRT systems are often based on one-to-one comparisons. The acceptable levels of performance and demographic differences for these systems may significantly differ by use case.

[20] GOV'T ACCOUNTABILITY OFF., PUB. NO. GAO-21-518 FACIAL RECOGNITION TECHNOLOGY: FEDERAL LAW ENFORCEMENT AGENCIES SHOULD BETTER ASSESS PRIVACY AND OTHER RISKS 20, https://www.gao.gov/assets/gao-21-518.pdf.

[21] *Id.* at 19.

[22] *Id.* at 19–20; *see also* DHS Directive 026-11, Use of Face Recognition and Face Capture Technologies, https://www.dhs.gov/sites/default/files/2023-09/23_0913_mgmt_026-11-use-face-recognition-face-capture-technologies.pdf.

One-to-many FRT is used by law enforcement agencies to identify or match people in images and video. These systems usually return a fixed number of candidate matches or candidate matches above a threshold similarity score. The results may be ordered by score and may show a numerical score or a score category.

While there is not comprehensive public data on FRT use by law enforcement nationwide, surveys indicate that FRT is widely used by federal, state, and local agencies.[23] A number of agencies operate their own FRT systems, based on photos from driver's licenses, arrests, and other government interactions. Agencies commonly have access to FRT systems maintained by other agencies, such as FBI's Next Generation Identification-Interstate Photo System (NGI-IPS), which can in turn incorporate results from other agencies. NGI-IPS, for example, incorporates results from 17 state agencies and two federal agencies and encompasses over 67 million arrest photos.[24] Commercial vendors also offer FRT services to law enforcement agencies which, as discussed further below, can heighten privacy impacts.

At the federal level, law enforcement agencies use FRT in support of their missions and pursuant to applicable policies. The FBI, for instance, uses one-to-many FRT to help identify perpetrators, victims, and witnesses as part of authorized investigations of criminal offenses.[25] The FBI also uses FRT to help identify and locate missing persons or other at-risk individuals, such as abducted children, or victims of child sexual abuse or human trafficking, and to identify deceased or incapacitated individuals. These uses of FRT can be faster and more efficient than other investigative methods, and they can provide unique leads that may not have been available or may have been impractical to obtain through other avenues. The FBI also uses FRT to structure and organize large volumes of lawfully obtained photo or video data, allowing investigators to more efficiently interpret the collected data.[26] Under the Department of Justice interim FRT policy, uses of FRT must be lawful and consistent with other DOJ policies. Among other requirements, FRT results alone may not be relied upon as the sole proof of a person's identity; activity protected by the First Amendment may not be the sole basis for using FRT; and personnel who use or approve FRT systems must receive relevant training; among other requirements.[27]

---

[23] *See* GAO-21-518 *supra* note 20 (finding that, in a GAO survey of 42 federal agencies with law enforcement responsibilities, 20 used FRT between 2015 and 2020); CLARE GARVIE, ALVARO BEDOYA & JONATHAN FRANKLE, THE PERPETUAL LINE-UP (2016) 93, 97 https://www.perpetuallineup.org/sites/default/files/2016-12/The%20Perpetual%20Line-Up%20-%20Center%20on%20Privacy%20and%20Technology%20at%20Georgetown%20Law%20-%20121616.pdf (reporting that, based on records requests to 106 state and local law enforcement agencies, at least 53 used, previously used, or planned to use FRT). There is limited data available about FRT use by Tribal law enforcement agencies.

[24] U.S. Dep't Just., Written Testimony in Connection with the United States Commission on Civil Rights' Examination of Civil Rights Implications of the Federal Use of Facial Recognition Technology (Mar. 21, 2024).

[25] The FBI's use of FRT is governed by the Department of Justice's interim FRT policy, which, among other requirements, mandates that FRT results alone may not be relied upon as sole proof of identity. Rather, an individual's identity must be confirmed through other analysis and/or investigation.

[26] This use case is a variation of 1:N FRT, where the gallery of images is drawn from the collected evidence in an investigation rather than from an established repository. The use case supports organizing and triaging media that has been collected in an investigation, which may be voluminous and unorganized.

[27] U.S. Dep't Just., Written Testimony in Connection with the United States Commission on Civil Rights' Examination of Civil Rights Implications of the Federal Use of Facial Recognition Technology (Mar. 21, 2024).

State and local law enforcement agencies similarly use FRT to support investigations. There is wide variation in applicable state laws, local ordinances, and law enforcement agency policies.[28] Differences include whether and when FRT may be used, protections for expressive activities, quality reviews of results from FRT, how the results from FRT may be used, which probe images may be used, which databases may be searched, what training is necessary, what information is recorded and audited, what information must be disclosed in discovery, and what public transparency must be provided. At one end of the spectrum, some agencies may use FRT under generally applicable laws and policies, but without a law or policy specific to FRT. At the other end, some jurisdictions have entirely prohibited law enforcement agencies from using FRT.

c. Iris Scanning

Iris scanning examines the unique tissue patterns in the donut-shaped part of an eye surrounding the pupil. Iris patterns do not appear to meaningfully change over time, and are protected by the cornea, limiting the potential for damage or mutilation.[29] Iris as a biometric modality is relatively new compared with FRT and other biometric modalities, with national-level matching capabilities coming online at the FBI in just the last 5 years.[30] The FBI's NGI Iris Service has over 3 million sets of iris images from over 2 million people.[31]

Iris scans are well-suited for identity confirmation in custodial settings, because they are highly accurate when collected properly and can be taken either from a short stand-off distance or without removing handcuffs. Iris scans can also be an effective supplement to other identification modalities for immigration and border screening.

Focusing on FRT for Identification in Law Enforcement Investigations

While FRT use by law enforcement agencies has significant benefits in developing leads, it also poses challenges for responsible use and governance of technology in criminal justice.

FRT poses significant privacy concerns because of the quantity, and likely long retention period, of data required for the system to be effective. FRT enables identifying people without

[28] *See generally* Mailyn Fidler & Justin (Gus) Hurwitz, *An Overview of Facial Recognition Technology Regulation in the United States*, in CAMBRIDGE HANDBOOK OF FACIAL RECOGNITION IN THE MODERN STATE (Rita Matulionyte & Monika Zalieriute, eds., 2024), https://doi.org/10.1017/9781009321211.018 ; GARVIE ET AL., *supra* note 23 at 121–50; Jameson Spivack & Clare Garvie, *A Taxonomy of Legislative Approaches to Face Recognition in the United States*, in REGULATING BIOMETRICS: GLOBAL APPROACHES AND OPEN QUESTIONS (Amba Kak ed., 2023), https://ainowinstitute.org/wp-content/uploads/2023/09/regulatingbiometrics-spivack-garvie.pdf ; JAKE LAPERRUQUE, CEN. DEM. & TECH., LIMITING FACE RECOGNITION SURVEILLANCE: PROGRESS AND PATHS FORWARD, (Aug. 23, 2022), https://cdt.org/insights/limiting-face-recognition-surveillance-progress-and-paths-forward/.

[29] *Iris Recognition*, NEC (Sept. 22, 2021), https://www.nec.com/en/global/solutions/biometrics/iris/index.html ("A person's iris pattern is unique and remains unchanged throughout life. Also, covered by the cornea, the iris is well protected from damage, making it a suitable body part for biometric authentication.").

[30] *The Eyes Have It: Iris Biometric Added to Next Generation Identification System* (Dec. 11, 2020), *https*://www.fbi.gov/news/stories/fbi-adds-iris-biometric-to-next-generation-identification-system-121120.

[31] FBI, FY 2025 Budget Request *supra* note 13 at 57 ("As of November 30, 2023, the NGI Iris Service consists of over 3.3 million sets of iris images representing more than 2.6 million unique identities.").

interacting with them or an object on which they left their fingerprints or DNA. The scale of FRT databases can be large, and the effort and cost of running FRT searches can be low.

Civil liberties are another area of concern. For instance, FRT could be misused to enable identification of people engaged solely in protected expressive activity. Going back to the Founding era, the United States has a rich tradition of anonymous civic discourse and protest, where privacy facilitates the expression of ideas and the assembly of groups. As the Supreme Court has recognized, there is a "vital relationship between freedom to associate and privacy in one's associations."[32]

Civil rights are another significant issue, in part due to possible biases in FRT systems and how they are used. For example, as noted above, public reporting indicates that there have been seven documented instances of mistaken arrests associated with the use of facial recognition technology, almost all involving Black individuals.[33] Many FRT systems deployed in the United States have higher false match (i.e., false positive) rates when applied to racial minorities, including people who are Black, Native American, Asian American, and Pacific Islanders.[34] Research has also demonstrated that FRT systems tend to perform worse on women, children, and the elderly, and some FRT algorithms used in the United States have biases also associated with eyewear, hairstyle, and other attributes.[35] Testing by the National Institute of Standards and Technology (NIST), discussed further below, indicates that in the last five years, some FRT developers have made significant progress in addressing differences in performance associated with demographics. Low absolute false match rates and lower relative rates—at least in the controlled settings of NIST's testing of recent algorithms—now exist across demographics, which include gender, age, and race.

The data used to train FRT systems is an important contributing factor. An FRT system generally performs best on faces that are similar to the faces used when training the system. Race, gender, age, and other attributes are often readily apparent from faces, unlike fingerprints and irises, which compounds the risk that these types of biometric systems will reflect demographic biases from training data. Research has demonstrated an "other-race" effect in FRT systems where, for example, systems built in the United States can perform better on white faces and systems

---

[32] *NAACP v. Alabama ex rel. Patterson*, 357 U.S. 449, 462 (1958).
[33] NAS 2024 Report *supra* note 4.
[34] NAS 2024 Report *supra* note 4; *Face Technology Evaluation – FTRE/FATE*, NAT'L INST. STANDARDS & TECH, https://www.nist.gov/programs-projects/face-technology-evaluations-frtefate.
[35] *E.g.*, NAS 2024 Report *supra* note 4; NIST *supra* note 34; Cynthia M. Cook et al., *Demographic Effects in Facial Recognition and Their Dependence on Image Acquisition: An Evaluation of Eleven Commercial Systems*, 1 IEEE TRANSACTIONS ON BIOMETRICS, BEHAV., & IDENTITY SCI. 32, 32-41 (2019), https://ieeexplore.ieee.org/document/8636231; Pawel Drozdowski et al., *Demographic Bias in Biometrics: A Survey on an Emerging Challenge*, 1 IEEE TRANSACTIONS ON TECH. & SOC'Y 89 (2020), https://ieeexplore.ieee.org/document/9086771; Philipp Terhörst et al., *A Comprehensive Study on Face Recognition Biases Beyond Demographics*, 3 IEEE TRANSACTIONS ON TECH. & SOC'Y. 16, 16-30 (2022), https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9534882. Related research has demonstrated similar disparities in other computer vision applications, such as gender classification. *E.g.*, Joy Buolamwini & Timnit Gebru, *Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification*, 81 PROCEEDINGS MACH. LEARNING RSCH. 1, 1 (2018), https://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf.

developed in Asia can perform better on Asian faces.[36] Balancing the demographics of training datasets can be a valuable step in improving performance disparities, though additional possible sources of bias remain in both FRT systems and how they are used.[37]

Agencies considering use of FRT must grapple with difficult policy decisions. Agencies should develop and enforce policies regarding the use of FRT and set clear rules for how and when the technology may be used, including guardrails that protect civil rights and liberties. Policies should, among other things, be transparent, be reflected in public documentation to the extent possible, include requirements for evaluating FRT uses, and provide for ongoing monitoring and mitigation of risks. In particular, policies should address the topics and uses described in greater detail below.[38]

a. Algorithm Evaluation

Before a law enforcement agency begins using an FRT system, it is essential to understand whether the benefits and risks of the system are appropriate for the intended use. Evaluating the algorithms in an FRT system is an important step and includes quantifying how well the system correctly matches faces when a match exists in the database, how well it rejects incorrect matches, and how these types of performance differ across demographic groups.[39] An FRT system with better performance can provide greater value in investigations and reduce the likelihood of harmful consequences.

NIST's ongoing Face Recognition Technology Evaluation (FRTE) program provides valuable algorithm performance benchmarks. The program includes one-to-many performance testing on over 350 FRT algorithms and one-to-one demographic testing (race, gender, and age) on over 500 algorithms.[40] The results of NIST's testing can help law enforcement agencies understand and compare algorithms.

As valuable as it is, the NIST testing program has important limitations. The datasets used in testing predominantly consist of images from a controlled or semi-controlled environment, where the subject is close to the camera, the subject is looking at or near the camera, and lighting is adequate. These datasets may be sufficiently representative for some criminal justice

---

[36] *See* PATRICK GROTHER ET AL., NAT'L INST. STANDARDS & TECH, FACE RECOGNITION VENDOR TEST (FRVT) PART 3: DEMOGRAPHIC EFFECTS (2019) https://doi.org/10.6028/NIST.IR.8280; P. Jonathan Phillips et al., *An Other-race Effect for Face Recognition Algorithms*, 8 ACM TRANSACTIONS ON APPLIED PERCEPTION 1 (2011), https://dl.acm.org/doi/10.1145/1870076.1870082; NIST *supra* note 34.

[37] *See* Valeriia Cherepanova et al., *A Deep Dive into Dataset Imbalance and Bias in Face Identification*, AIES (2023) https://dl.acm.org/doi/fullHtml/10.1145/3600211.3604691.

[38] Law enforcement agencies can use existing templates to guide policy development, including the FRT Policy Development Template. *Face Recognition Policy Development Template*, BUREAU JUST. ASSISTANCE (Dec. 2017), https://bja.ojp.gov/sites/g/files/xyckuh186/files/Publications/Face-Recognition-Policy-Development-Template-508-compliant.pdf.

[39] In NIST's 1:N FRTE program, the core metrics are the false negative identification rate and the false positive identification rate. There are many other possible performance and bias metrics to compute for FRT systems, beyond those included in NIST's valuable program, like for other AI systems. There are also other types of bias to consider, such as people with disabilities. Testing with other metrics and for other types of biases may be appropriate depending on the intended use.

[40] *See* NIST *supra* note 34. As NIST explains in its demographic testing methods, it is possible for many FRT algorithms to extrapolate one-to-many demographic results from one-to-one results.

applications where test results may be indicative of the system's performance, such as FRT use to match booking photos to driver's license photos.

The datasets used in NIST's current testing may, however, not be representative of law enforcement investigative uses of FRT. These uses often involve images that are not from controlled environments, such as still images from surveillance camera footage. The subject may be distant, looking away, dimly or unevenly lit, or located below the camera. The camera may be of low quality or introduce distortions, and the image or the subject's face may be occluded. These significant differences in context make it difficult to generalize the results of NIST's current testing to FRT performance and biases when used in law enforcement investigations.

NIST is in the process of reestablishing a line of performance testing for FRT use on images from videos, including from surveillance cameras. This type of testing, the Face In Video Evaluation (FIVE), may be more representative of law enforcement investigation settings.[41]

Law enforcement agencies considering use of FRT should establish testing requirements for performance and biases. This testing should, to the extent possible, be representative of real-world deployment contexts and follow standardized methodologies. NIST's testing program is an important starting point, and the ISO/IEC 19795 standards on biometric performance testing also provide valuable guidance. Agency policies should specify the nature of and benchmarks for testing and should ensure retention and disclosure of testing results to the extent feasible. Policies should also require that that vendors provide evaluation results for the system version that is procured by agencies, not results for a previous or adjusted versions of the system, and that the documentation and results from vendors be sufficient to allow for independent evaluation and/or auditing.

Continuous monitoring, discussed further below, can provide additional information about the real-world performance and biases of FRT systems. When pre-deployment testing is not fully representative of how an FRT system will be used, post-deployment monitoring is especially important.

At the federal level, OMB Memorandum M-24-10 identifies law investigative uses of FRT as AI use cases that presumptively require heightened risk management practices, including performance and bias testing in real-world conditions. OMB Memorandum M-24-18 further directs federal agencies that procure FRT capabilities to ensure that they have been tested by NIST, where practicable. State, local, Tribal, and territorial law enforcement agencies should also implement these practices, and federal grantmaking agencies should require these practices when providing financial support for the procurement or use of FRT, accounting for the differing missions and resources of grant recipients.[42]

[41] FIVE will report performance measurements for FRT algorithms but, importantly, may not include measurements of demographic differences.

[42] For example, in some instances, it may be appropriate for a state, local, Tribal, or territorial law enforcement agency to evaluate a prospective FRT use on the basis of real-world FRT testing conducted by, on behalf of, or in coordination with other law enforcement agencies, provided that the testing is representative of the agency's FRT uses.

b. Database Selection

Agencies may have access to multiple FRT systems with different databases, and some FRT systems have the capability to run searches against multiple databases. Depending on the agency and use case, searches can be run against criminal records, driver's license photographs, other agencies' databases, or commercial databases.

While the likelihood of developing a useful investigative lead generally increases with an expanded search, so too does the likelihood of an FRT system returning candidate matches with high similarity scores—and possibly high visual similarity for human reviewers—that are not the person to be identified. A larger set of databases or data can also increase the risk that a user may unintentionally exceed their authority by searching a dataset or accessing a photo for an unauthorized purpose.

Agency policies should clearly articulate to users what datasets are available for each type of search, as well as the purposes for which a search of a given dataset is authorized. Agency agreements to access external data sources should also contain appropriate restrictions on the use of the data being accessed. Agencies should establish processes to log FRT uses, enabling auditing to ensure that searches have been conducted for authorized purposes.

In selecting databases to use, agencies should give careful consideration to how the underlying data was collected. Some commercial FRT services make use of databases that contain millions, or even billions of images scraped from social media and other online services and websites. This repurposing of personal photographs, in a context different from the ones in which they were originally created and shared—potentially without consent and contrary to expectations—potentially raises questions of law, policy, and ethics. These types of commercial services also generally have less reliable, less complete, and less current information associated with photos than do FRT systems based on government identification records, which can misdirect investigative efforts. Some law enforcement agencies prohibit the use of these types of systems, and others permit their use only in certain types of investigations. OMB Memoranda M-24-10 and M-24-18 specifically direct federal agencies to carefully consider whether and when use of these types of FRT systems is appropriate.

Law enforcement agencies should not use FRT systems trained on photos or built with other information that was collected in violation of laws, federal government guidance, or agency policy. Agencies should also specifically articulate the authority that permits the collection of FRT biometric data or associated personally identifiable information, which should be reflected in public documentation whenever possible.

c. Use of Facial Recognition

Law enforcement agencies substantially differ in their policies regarding when an investigation may make use of facial recognition. At some agencies, facial recognition is available for all criminal investigations. At others, only certain types of investigations may make use of FRT, such as for violent crimes and child safety. As noted above, at some law enforcement agencies, FRT use is prohibited.

There is further divergence in the predication standards that law enforcement agencies implement for using facial recognition. At some agencies, investigators have discretion about turning to FRT. At others, investigators must meet a reasonable suspicion standard. And at some agencies, investigators must have probable cause to conduct an FRT search.

Real-time use of FRT is another area where law enforcement agencies have differing policies. Some prohibit it, while others allow it but restrict use directed at protected speech activities.

Law enforcement agencies that use FRT should establish policies that clearly specify when FRT may be used. These policies should be public and easily accessible to the greatest extent possible.

Agency policies should describe the types of investigations in which FRT use is appropriate, taking into account factors such as the type of the criminal offense, the likelihood of generating a true match, the evaluated performance of the FRT system, and the quality of relevant data. Agency policies should also delineate the circumstances in which it is appropriate to conduct FRT searches, such as to provide a lead for identifying a witness, perpetrator, victim, or a person who is missing or otherwise believed to be at risk of harm. Policies should additionally specify what predication is necessary to take the step of an FRT search.

A policy should also describe the level of supervisory review, if any, necessary before conducting a search. Policies should also distinguish between the different types of FRT use and account for the varying levels of risk of harm. For example, using FRT to identify an unknown perpetrator need not be treated the same as using FRT to organize and triage collected media.

In addition to specifying when FRT may be used, policies should specify when it may not be used. At minimum, policies should prohibit the use of FRT solely based on constitutionally protected activities (e.g., a First Amendment-protected protest), to facilitate unlawful discrimination, or in any other way that would be inconsistent with legal requirements or other policies.

d. Image Quality

The quality of the probe image submitted to an FRT system is a critical factor in whether the system can return a match. Many factors can affect the performance and biases of an FRT system, including photo resolution and clarity; the subject's pose and attire; lighting; occlusions; and the camera's position, sensor, and lens. Additionally, modifications to the probe image—including to size, aspect ratio, or coloration—could potentially have an adverse impact on FRT results. In general, searches with lower-quality images are less likely to return matches. In a law enforcement investigation, however, only lower-quality images may be available.

Law enforcement agencies can mitigate risks associated with probe photo selection by establishing minimum quality criteria.[43] Where possible, these criteria should be set by an

[43] The Facial Identification Scientific Working Group (FISWG)—a consortium of state, local, federal, and international law enforcement agencies as well as FRT vendors and academics—provides voluntary image quality

independent entity with expertise in FRT, such as a testing or standards-setting organization. These criteria may differ depending on the type of use, the feasibility of alternative investigative steps, and agency resources. An agency might conclude, for example, that the minimum quality criteria for locating a missing person or identifying a victim of child sex trafficking should be different from the criteria for identifying a witness to a nonviolent crime.

The policies that law enforcement agencies establish for FRT should also address use of probe images that are not photographs of the person to be identified. These images may include sketch drawings, generated images, or images of people who are "lookalikes" for the subject. Searches with these types of images can be more prone to incorrect matches,[44] so a policy should establish when (if ever) they are permitted and should require heightened safeguards when they are used.

Law enforcement agencies should also implement policies that prohibit the use of FRT with probe images that were collected in violation of law or another applicable policy.

e. Quality Control for Results

The FRT systems used in law enforcement investigations are inexact, and a system may return candidate matches that are not the subject.[45] Law enforcement agencies should apply a minimum similarity threshold for candidate results, which may vary depending on the nature of the investigation and should only be overridden in exigent circumstances.

Human review is also essential. Law enforcement agencies should require that an examiner who is trained to compare faces and mitigate bias—and who ideally is independent of the case team—manually reviews results before they are used in an investigation. When there are multiple candidate results from an FRT system, an examiner should review all top results, and similarly a case team should consider all candidates returned by an examiner before focusing on one candidate result for further investigation.

---

standards. *Image Factors to Consider in Facial Image Comparison*, FACIAL IDENTIFICATION SCI. WORKING GRP. (May 28, 2021), https://fiswg.org/fiswg_image_factors_to_consider_in_facial_img_comparison_v1.0_2021.05.28.pdf. The International Standardization Organization has worked with other organizations to develop the ISO/IEC 30137 series, which outlines effective video system performance for FRT and other uses. *See generally ISO/IEC JTC 1/SC 37: Biometrics*, INT'L STANDARDIZATION ORG. & INT'L ELECTROTECHNICAL COMM'N, https://www.iso.org/committee/313770.html. *ISO/IEC 30137-1:2024: Information technology — Use of biometrics in video surveillance systems*, INT'L STANDARDIZATION ORG. & INT'L ELECTROTECHNICAL COMM'N (2024), https://www.iso.org/standard/87734.html.

[44] *See, e.g.*, NAT'L INST. STANDARDS & TECH, NIST INTERAGENCY REPORT 8009, FACE RECOGNITION VENDOR TEST 4 (2014), https://nvlpubs.nist.gov/nistpubs/ir/2014/NIST.IR.8009.pdf (documenting high error rates in sketch recognition); Christian Galea & Reuben A. Farrugia, *Forensic Face Photo-Sketch Recognition Using a Deep Learning-Based Architecture*, 24 IEEE SIGNAL PROCESSING LETTERS 1586 (2017), https://ieeexplore.ieee.org/document/8025793 (discussing performance deep learning for face photo-sketch recognition); CLARE GARVIE, GARBAGE IN, GARBAGE OUT (2019), https://www.flawedfacedata.com/ ("Even the most detailed sketches make poor face recognition probe images. . . . The most likely outcome of using a forensic sketch as a probe photo is that the system fails to find a match—even when the suspect is in the photo database available to law enforcement.").

[45] *See* NAS 2024 Report *supra* note 4, at 1,6.

Law enforcement agencies should take steps to minimize the risks of automation bias (i.e., examiner deference to system output) and confirmation bias (i.e., reinforcing an examiner's beliefs about a subject or the system). These steps could include removing similarity scores or ranking information from the results shown to examiners, reminding examiners of the limitations of FRT systems, and providing appropriate training on facial comparison and mitigating biases (discussed further below).

f. Uses of FRT Results

In addition to specifying when and how investigators can use FRT, law enforcement agency policies should establish permissible uses of results. Because of the limitations of FRT, policies should specify that FRT search results should be considered a lead and not sufficient to establish probable cause or a positive identification without corroboration. Policies should describe when and how FRT can support probable cause. This can be a complex issue where FRT plays a role in witness identification, which may also involve comparing faces. At least one law enforcement agency prohibits conducting a lineup based solely on an FRT investigative lead without independent and reliable evidence linking the suspect to the crime.[46]

g. End-to-End Evaluation and Continuous Monitoring

When law enforcement agencies use FRT systems in investigations, it is essential to understand the technical performance and biases of these systems, as noted above. It is also important to understand the broader context for and impacts of FRT use, including why investigators use it, how it affects investigations, and how it affects the people who appear in results. This type of "end-to-end" operational evaluation is a best practice for AI governance and, at the federal level, encouraged by OMB Memorandum M-24-10.

Law enforcement agencies should consider implementing end-to-end evaluation for uses of FRT. This evaluation could address questions like: What are the types of cases where investigators turn to FRT, and why do they use FRT instead of other investigative methods? How valuable are the leads from FRT in advancing investigations? How often does FRT generate a lead for investigators that could not have been developed otherwise, or that would have taken considerably more time or resources otherwise? How often does FRT lead investigators to focus on a person who is later determined to not be relevant to an investigation? Answering basic questions like these can be important for evaluating the benefits and risks of FRT use, and can help reinforce community trust by demonstrating the practical impacts of FRT.

Continuous monitoring is another form of evaluation that law enforcement agencies should consider. An FRT system's behavior is affected by both probe and gallery photos. If there are changes in either type of photo—for example, if a law enforcement agency starts focusing on a particular type of photo as evidence or if a state changes its driver's license photo format—that can impact the performance and biases of the system. FRT vendors also update their algorithms, which can also affect performance and biases. Continuously keeping track of real-world

[46] U.S. COMM'N CIV. RTS., THE CIVIL RIGHTS IMPLICATIONS OF THE FEDERAL USE OF FACIAL RECOGNITION TECHNOLOGY 112-13 (2024), https://www.usccr.gov/files/2024-09/civil-rights-implications-of-frt_0.pdf (Stmt. of Vice Chair Nourse) (discussing Detroit Police Department practices and policies).

performance and bias statistics can substantiate a system's ongoing value and can alert an agency if there are changes that need attention. OMB Memorandum M-24-10 requires federal law enforcement agencies to implement continuous monitoring for FRT uses.

h. Restrictions on FRT Use and Logging

Policies should prohibit the use of FRT systems unless the agency has approved the system, the user, and the use. Policies and procedures should prohibit and establish consequences for unauthorized or improper use of a FRT biometric system (including examples discussed above, such as use based solely on constitutionally protected activity). Agencies should retain detailed internal logs of FRT system use for auditing and ensuring compliance with requirements.

i. Transparency Regarding FRT Use

As feasible, law enforcement agencies should adopt policies that require them to publicly disclose their use of FRT use, including details of the system in use, and the nature and purpose of the use. Agencies should also engage with community stakeholders about FRT and, to the extent possible, provide transparent responses about how they use FRT.

At the federal level, OMB Memorandum M-24-10 requires these practices for law enforcement uses of FRT. Federal law enforcement agencies must provide public transparency about FRT uses in an annual AI inventory. Agencies are also required to engage with stakeholders to obtain their input.

j. Data Management

Law enforcement agency policies on FRT should describe the collection, management, storage, and retention requirements for requests, probe images, and results. Policies should also describe security, privacy, recordkeeping, and audit requirements. Agencies should ensure compliance with any connected system policies, including FBI's Criminal Justice Information Services (CJIS) Security Policy.[47]

k. Training

Documented and effective training is critical for the successful implementation of FRT. A policy should establish training requirements for personnel who will interact with FRT or its results. Training should address all aspects of an agency's FRT policy, including when investigators may run an FRT search and how the results may be used. Training should also provide background on the technology and its limitations, including how incorrect matches can occur and may disproportionately affect certain demographic groups.

[47] *See generally* BUREAU JUST. ASSISTANCE, FACE RECOGNITION POLICY DEVELOPMENT TEMPLATE (Dec. 2017), https://bja.ojp.gov/sites/g/files/xyckuh186/files/Publications/Face-Recognition-Policy-Development-Template-508-compliant.pdf (discussing best practices and applicable standards for security).

Training related to FRT should also cover the risks of human biases, such as automation and confirmation bias, when using FRT.[48] These biases can cause personnel to place undue weight on certain results, and strategies are available to mitigate the risk.[49] Demographic bias (explicit or implicit) can also affect human judgment and should be addressed in training. Research has shown that the innate ability to recognize faces varies widely and that people less reliably identify others from a different race.[50] This bias may compound with biases in FRT algorithms.[51]

Training for FRT use in law enforcement should be appropriate to a person's role and should convey the information necessary for responsibly submitting a probe image, analyzing results from an FRT system, and using the results in law enforcement activities. Roles for training that may be common across law enforcement agencies include:[52]

- Facial Examiner: Compares a probe photo to candidate matches from an FRT system to develop possible investigative leads.
- Collector (Includes Investigators): Obtains probe images for use with an FRT system.
- Facial Reviewer (Includes Investigators): Reviews results of an FRT search adjudicated by a facial examiner.
- Supervisor: Oversees personnel involved with FRT and ensures compliance with law and policy.

Failure to properly train individuals who interact with FRT systems can increase the risk of potential errors at each step of the facial recognition process, which could ultimately impact individuals' privacy, civil rights, and civil liberties.

Automated License Plate Recognition

License Plate Readers (LPRs) are cameras with computer vision capabilities designed to detect and capture information from license plates within their field of view. Computer vision is an area of AI that can identify patterns and objects in images. In the criminal justice context, LPR cameras can be mounted on patrol vehicles to identify vehicles in connection with criminal investigations, including stolen vehicles, vehicles owned by wanted persons, vehicles involved in

[48] *See* Reva Schwartz et al., NAT. INST. STANDARDS & TECH. SPECIAL PUBL'N 1270, TOWARDS A STANDARD FOR IDENTIFYING AND MANAGING BIAS IN ARTIFICIAL INTELLIGENCE 26 (2022), https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.1270.pdf. (discussing systemic biases, statistical and computational biases, and human biases).

[49] *See* Samuel Peterson et al., RAND, *Finding a Broadly Practical Approach for Regulating the Use of Facial Recognition by Law Enforcement* 32 (Feb. 15, 2023), https://www.rand.org/pubs/research_reports/RRA2249-1.html (discussing sources of bias, including human decisionmaking).

[50] Jacqueline G. Cavazos et al., *Learning Context and the Other-Race Effect: Strategies for Improving Face Recognition*, 157 VISION RSCH., Apr. 2019, 169, 169–83, https://nij.ojp.gov/library/publications/learning-context-and-other-race-effect-strategies-improving-face-recognition.

[51] *See* Peterson et al., *supra*, note 49.

[52] *See* FACIAL IDENTIFICATION SCI. WORKING GRP., GUIDE FOR ROLE-BASED TRAINING IN FACIAL COMPARISON (2020), https://fiswg.org/fiswg_guide_for_role-based_training_in_facial_comparison_v1.0_20200717.pdf (listing different roles in a facial comparison environment including assessor, reviewer, examiner, manager, supervisor, collector, technical reviewer, and trainer).

an ongoing emergency like kidnapping, or vehicles being operated in an unlawful manner, enabling timely law enforcement action. LPR data also enables retrospectively tracking the movements of a vehicle of interest.

LPR cameras can be mounted in fixed locations to identify vehicles entering or exiting sensitive locations, including prisons or other government facilities, as well as known locations of criminal activity to identify potential criminals or enable facility security activities.

The collected license plate numbers can be cross-referenced against law enforcement databases to identify vehicles. That information can be used to identify vehicles that may require law enforcement action, such as stolen vehicles or suspect vehicles.[53] LPR cameras, including both fixed and vehicle mounted cameras, can assist in locating vehicles during Amber alerts, Ashanti alerts, silver alerts, or similar emergency situations.[54]

Some LPR systems are operated by law enforcement agencies, using their own cameras typically on their own patrol vehicles. Other LPR systems are commercial services, with networks of participating cameras (e.g., at parking lots) that agencies can subscribe to.

LPR systems are in widespread use. Almost all large local law enforcement agencies have an LPR program, as do many smaller agencies.[55] Law enforcement agencies collect billions of LPR records per year.

Like biometrics used for identification, LPR systems can serve as an effective tool for monitoring vehicles in connection with criminal investigations. LPR successes include apprehending violent offenders and rescuing abducted children.[56] The use of LPR systems, however, comes with some risks. For example, LPR systems can misread plates or misidentify stolen vehicles.[57] In addition, as with all large stores of data, agencies may create privacy risks if they do not take the necessary steps to properly secure LPR data and dispose of it after it is no longer needed.

LPR systems, like FRT systems, must be accompanied by strong policy and procedural guardrails to ensure appropriate use.

---

[53] LPR systems in the United States are operated by both law enforcement agencies as well as commercial providers who operate the system as a service. As a result, the database referenced and system data retention policies are specific to each LPR system.

[54] *See* U.S. DEP'T JUST, BUREAU JUST. ASSISTANCE, FACT SHEET: NATIONAL ASHANTI ALERT NETWORK (2021). https://bja.ojp.gov/sites/g/files/xyckuh186/files/media/document/National-Ashanti-Alert-Network-Fact-Sheet.pdf ("Ashanti Alerts, once implemented, can provide rapid dissemination of information to law enforcement agencies, media, and the public about adults who have been reported missing, along with suspect information in cases of abduction.").

[55] U.S. DEP'T JUST., OFF. JUST. PROGS., BUREAU JUST. STATS., LOCAL POLICE DEPARTMENTS, 2013: EQUIPMENT AND TECHNOLOGY (2015), https://bjs.ojp.gov/content/pub/pdf/lpd13et.pdf ("An estimated 17% (about 2,000) of departments used automated license plate readers in 2013. This total included more than three-quarters of the departments serving 100,000 or more residents.").

[56] ANGEL DIAZ & RACHEL LEVINSON-WALDMAN, AUTOMATIC LICENSE PLATE READERS: LEGAL STATUS AND POLICY RECOMMENDATIONS FOR LAW ENFORCEMENT USE, BRENNAN CTR. (2020), https://www.brennancenter.org/our-work/research-reports/automatic-license-plate-readers-legal-status-and-policy-recommendations.

[57] *Id.*

LPR Risks and Mitigation

a. Procedural Risks

Agencies should establish clear policies that outline where, when, and for what purpose an LPR camera can be placed to enhance criminal justice operations. Policies and procedures should also clearly address retention periods for LPR records, as well as the circumstances, if any, in which they can be searched.

b. Accuracy and Reliability

Accuracy in LPR technology is critical to avoid false positives, misidentifications, and misdirected law enforcement actions. Agencies should prioritize data quality, with standards and procedures for identifying and correcting errors.

c. Safeguarding Privacy and Data Security

Protecting the security of LPR data is a priority. This includes implementing robust data encryption, access restrictions, and audit mechanisms to prevent unauthorized access and misuse. Agencies should adopt strong data security protocols to reinforce public confidence in LPR technology while safeguarding sensitive information.

d. LPR Data Sharing

Agencies should establish data sharing protocols if LPR data is shared with other agencies or third-party organizations. Data sharing agreements should include safeguards to protect privacy, and ensure all parties uphold the same standards of data security and ethical use.

e. Training

Proper training for using LPR systems is essential to ensure law enforcement personnel are aware of legal, ethical, and operational standards. This training should cover privacy protections, the importance of data accuracy, and the need to prevent bias in the use of LPR technology.

III. Forensic Analysis

Introduction

Forensic analysis of physical, digital, and multimedia evidence is central to criminal investigation and litigation. In recent years, uses of AI—including statistical techniques and machine learning[1]—have accelerated the trend in forensic disciplines from subjective judgment in analyzing and interpreting forensic evidence toward more objective approaches.[2] This paradigm shift has potential to improve the reproducibility and accuracy of forensic methods, mitigate the possible human biases and examiner variation that may affect forensic analysis, and reinforce public trust in the criminal justice system.[3] AI may also provide new capabilities and reduce the time and cost of analysis processes.

Professionals in forensic science have a responsibility to rigorously validate methods of analysis, and that responsibility remains with uses of AI.[4] Properly designed validation studies can empirically demonstrate that a forensic method is reproducible and accurate, both in principle and

[1] As noted in the Introduction, definitions of AI vary substantially, with some broadly encompassing statistics and others emphasizing recent advances in machine learning. This chapter considers statistical and machine learning methods together because the opportunities and challenges for forensic analysis are broadly similar, and because the approach is consistent with the definition of artificial intelligence in OMB Memoranda M-24-10 and M-24-18. *See* R. SHUTE ET AL., WHAT FSSP LEADERS SHOULD KNOW ABOUT ARTIFICIAL INTELLIGENCE AND ITS APPLICATION TO FORENSIC SCIENCE 4 (Nat'l Inst. Just. 2023), https://forensiccoe.org/private/65cfa81c601c4 (noting that automated and semi-automated systems used in forensic analysis require similar assessment, regardless of whether they are categorized as AI).

[2] *See* EXECUTIVE OFFICE OF THE PRESIDENT, PRESIDENT'S COUNCIL OF ADVISORS ON SCIENCE AND TECHNOLOGY, FORENSIC SCIENCE IN CRIMINAL COURTS: ENSURING SCIENTIFIC VALIDITY OF FEATURE-COMPARISON METHODS 46-54 (2016) ("PCAST Report"), https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_forensic_science_report_final.pdf (describing differences between subjective and objective forensic methods and explaining that objective methods are generally preferable). *See also* NATIONAL RESEARCH COUNCIL, STRENGTHENING FORENSIC SCIENCE IN THE UNITED STATES: A PATH FORWARD (2009), https://nap.nationalacademies.org/catalog/12589/strengthening-forensic-science-in-the-united-states-a-path-forward; JOSÉ ALMIRALL ET AL., AM. ASS'N ADVANCEMENT SCI., FORENSIC SCIENCE ASSESSMENTS: A QUALITY AND GAP ANALYSIS: FIRE INVESTIGATION (2017), https://www.aaas.org/sites/default/files/s3fs-public/reports/Fire%2520Investigation_0.pdf; WILLIAM THOMPSON ET AL., AM. ASS'N ADVANCEMENT SCI., FORENSIC SCIENCE ASSESSMENTS: A QUALITY AND GAP ANALYSIS: LATENT FINGERPRINT EXAMINATION (2017), https://www.aaas.org/sites/default/files/s3fs-public/reports/Latent%2520Fingerprint%2520Report%2520FINAL%25209_14.pdf; Jonathan J. Koehler et al., *The Scientific Reinvention of Forensic Science*, 120 PROCS. NAT'L ACAD. SCIS. e2301840120 (2023), https://doi.org/10.1073/pnas.2301840120.

[3] *See, e.g.*, Mark Barash et al., *Machine Learning Applications in Forensic DNA Profiling: A Critical Review*, 69 FORENSIC SCI. INT.: GENETICS 13 (2024).

[4] *E.g.*, U.S. DEP'T JUST., CODE OF PROFESSIONAL RESPONSIBILITY FOR THE PRACTICE OF FORENSIC SCIENCE (2016) ("DOJ Code"), https://www.justice.gov/sites/default/files/code_of_professional_responsibility_for-the_practice_of_forensic_science_08242016.pdf ("5. Conduct research and forensic casework using the scientific method or agency best practices. Where validation tools are not known to exist or cannot be obtained, conduct internal or inter-laboratory validation tests in accordance with the quality management system in place."; "8. Conduct examinations that are fair, unbiased, and fit-for-purpose."; "10. Ensure interpretations, opinions, and conclusions are supported by sufficient data and minimize influences and biases for or against any party.").

as applied in a particular case.[5] Validation can also enable forensic practitioners to convey valuable context for the results of forensic analysis, such as the likelihood of results occurring by chance or reflecting errors.

The use of AI in forensic science can add complexity to validation. Models built from data can have nuanced performance characteristics and incorporate demographic biases. Models can also be sensitive to subtle differences between the data used during development and the data encountered in real-world use, such as differences in how data is collected and prepared for forensic analysis. The implementation of AI systems can be complicated and proprietary, which can make it difficult to validate that a system implements a forensic method as intended. These possible sources of additional complexity reinforce the importance in forensic science of responsible practices for developing, validating, and revalidating methods of analysis.

Explainability is also important in forensic science.[6] Practitioners have a responsibility to explain in a straightforward manner the data that they analyzed, the methods that they applied, and the interpretations, observations, and conclusions that resulted from applying the methods to the data.[7] AI models may not be readily understandable by humans and may learn from correlations in data that are difficult to discern and not necessarily causal.[8] Differences like these may affect how stakeholders in the criminal justice system are able to explain forensic processes that involve AI.

Expert oversight is another critical dimension of forensic science.[9] AI should be a complement to the expertise of forensic practitioners, such as by recommending next steps for human consideration, checking human analysis, or providing a basis on which an expert might

---

[5] *See* PCAST Report, *supra* note 2, at 42-43 (explaining the distinction between "foundational validity" and "validity as applied").

[6] The term "explainability" has a particular technical meaning in evaluating artificial intelligence systems. Here, this report uses it in the colloquial sense.

[7] Professional codes in forensic science often address explanation of data, methods, and conclusions. *E.g.*, DOJ Code, *supra* note 4 ("12. Prepare reports and testify using clear and straightforward terminology, clearly distinguishing data from interpretations, opinions, and conclusions. Reports should disclose known limitations that are necessary to understand the significance of the findings."; "15. Honestly communicate with all parties (the investigator, prosecutor, defense, and other expert witnesses) about all information relating to their analyses, when communications are permitted by law and agency practice."). *See* EXECUTIVE OFFICE OF THE PRESIDENT, NAT'L SCI. & TECH. COUNCIL, STRENGTHENING THE FORENSIC SCIENCES 24 (2014), https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/NSTC/forensic_science___may_2014.pdf (noting that, in a review of "more than 45 codes of ethics in use by various forensic science organizations," a commonality was "the need to . . . provide clear and objective testimony"). Courts may also interpret legal standards for the admissibility of expert evidence, discussed further below, to require forms of explanation. *E.g.*, *Zenith Elecs. Corp. v. WH-TV Broad. Corp.*, 395 F.3d 416, 419 (7th Cir. 2005) (explaining that "[a]n expert must offer good reason to think that his approach produces an accurate estimate using professional methods, and this estimate must be testable," such that "[s]omeone else using the same data and methods must be able to replicate the result," and that if an expert "could or would not explain how his conclusions met [these] requirements, he was not entitled to give expert testimony").

[8] Leo Breiman, *Statistical Modeling: The Two Cultures (with Comments and a Rejoinder by the Author)*, 16 STAT. SCI. 199-231 (2001), https://doi.org/10.1214/ss/1009213726.

[9] *E.g.*, DOJ Code, *supra* note 4 ("9. Make and retain contemporaneous, clear, complete, and accurate records of all examinations, tests, measurements, and conclusions, in sufficient detail to allow meaningful review and assessment by an independent professional proficient in the discipline.").

arrive at a conclusion.[10] Practitioners have an essential role in interpreting the output from AI systems, forming and explaining conclusions, and (when necessary and appropriate) offering expert testimony in litigation. Training for practitioners can also ensure that forensic methods involving AI are used properly and that results from analysis are accurately characterized.

The first part of this chapter describes how AI is currently being used in forensic analysis, as well as research suggesting areas of forensic science where AI may be used in the future. The breadth of use cases speaks to the immense potential of AI for forensic science. The second part of the chapter discusses challenges for AI use in forensic science and means of mitigating the risks associated with AI use. The chapter closes with a set of recommendations for forensic research and practice.

Current Uses of AI in Forensic Analysis

Several factors have limited the integration of AI into forensic analysis. These include the complex nature of forensic science, the standards for admissibility of forensic evidence in litigation, limited availability of high-quality and real (or realistic) data from forensic analysis settings, and resource constraints on forensic researchers and practitioners. As a result, real-world use of AI in forensics is presently limited outside of a few contexts.

As discussed in the Identification and Surveillance chapter, AI is widely used in biometric analysis, including fingerprint, palm print, iris, and face comparison. These methods complement other means of identification, and they are often used to retrieve candidate matches from a database for further forensic analysis or law enforcement investigation.

In DNA analysis, probabilistic genotyping can enable analysts to interpret complex samples that contain small amounts of DNA, mixed DNA, or damaged DNA that would be challenging for traditional analysis methods.[11] The statistical methods used for this type of forensic analysis typically estimate a likelihood ratio, comparing the probability of DNA observations given one proposition (e.g., that a defendant contributed genetic material to a sample) to the probability of observations given a stated alternative proposition (e.g., that an unknown person in a relevant population contributed to the sample).[12] The particular statistical analysis, population data, and

[10] H. Swofford & C. Champod, *Implementation of Algorithms in Pattern & Impression Evidence: A Responsible and Practical Roadmap,* 3 FORENSIC SCI. INT'L: SYNERGY (2021), https://doi.org/10.1016/j.fsisyn.2021.100142.

[11] *See* Michael D. Coble & Jo-Anne Bright, *Probabilistic Genotyping Software: An Overview*, 38 FORENSIC SCI. INT'L: GENETICS 219 (2019). As noted at the outset of the chapter, definitions of artificial intelligence differ. Some practitioners, for example, may not consider probabilistic genotyping to be a type of AI. *See* R. SHUTE ET AL., *supra* note 1, at 3 (describing disagreement about whether probabilistic genotyping is AI and noting it is at minimum part of an "automated system" that can exhibit bias).

[12] *See, e.g.*, Peter Gill et al., *DNA Commission of the International Society of Forensic Genetics: Recommendations on the Evaluation of STR Typing Results That May Include Drop-out and/or Drop-in Using Probabilistic Methods*, 6 FORENSIC SCI. INT'L: GENETICS 679 (2012), http://dx.doi.org/10.1016/j.fsigen.2012.06.002; Hannah Kelly et al., *A Description of the Likelihood Ratios in the Probabilistic Genotyping Software STRmix*, 2 WIREs FORENSIC SCIENCE (2020), https://doi.org/10.1002/wfs2.1377.

assumptions underlying probabilistic genotyping methods can vary, leading to different results.[13] The extent of validation research also differs by probabilistic genotyping method.[14]

Forensic genetic genealogy is another type of DNA analysis that can involve statistical models.[15] These methods of analysis can enable the generation of leads by comparing a DNA sample to a large database of samples and estimating possible genealogical relationships.

Narcotics tracing is another area of forensics where AI is in use today. The Drug Enforcement Administration (DEA), for example, uses machine learning models to classify the geographic region of origin for samples of heroin and cocaine.[16] The DEA's system was developed with authentic drug samples and can detect anomalies in analysis and low-confidence results. This program is valuable for understanding trends in drug trafficking, though the results are not presently used as evidence in court.

AI features are increasingly common in tools for forensic analysis of digital and multimedia evidence.[17] Computer vision techniques, for example, can assist forensic examiners in searching large and unorganized collections of photos and videos for specific content, such as weapons, nudity, or violence, that may be helpful to an investigation. Natural language processing methods can similarly help identify files and communications that relate to particular topics. Machine translation can support examiners who are analyzing evidence that involves multiple languages. AI can be valuable, in some circumstances, for identifying and analyzing evidence that may have been created or modified by AI (e.g., "deepfake" images, videos, and audio).[18] These uses of AI for analysis of digital and multimedia evidence are predominantly, for now, in support of investigative steps rather than expert conclusions offered in court.

---

[13] *See* John Buckleton et al., *A Diagnosis of the Primary Difference Between EuroForMix and STRmix*™, 69 J. FORENSIC SCI. 40 (2024), https://doi.org/10.1111/1556-4029.15387; Peter Gill et al., *A Review of Probabilistic Genotyping Systems: EuroForMix, DNAStatistX and STRmix*™, 12 GENES 1559 (2021), https://www.mdpi.com/2073-4425/12/10/1559; Susan A. Greenspoon et al., *A Tale of Two PG Systems: A Comparison of the Two Most Widely Used Probabilistic Genotyping Systems in the United States*, 69 J. FORENSIC SCI. 1840 (2024), https://doi.org/10.1111/1556-4029.15571.

[14] See the discussion below for further detail about validation of probabilistic genotyping software.

[15] *See* FORENSIC TECH. CTR. OF EXCELLENCE., AN INTRODUCTION TO FORENSIC GENETIC GENEALOGY TECHNOLOGY FOR FORENSIC SCIENCE SERVICE PROVIDERS (Nat'l Inst. of Just. 2022), https://forensiccoe.org/private/66291221e66ec; U.S. DEP'T JUST., INTERIM POLICY ON FORENSIC GENETIC GENEALOGICAL DNA ANALYSIS AND SEARCHING (2019), https://www.justice.gov/olp/page/file/1204386/dl.

[16] U.S. DEP'T JUST., 2023 AI USE CASE INVENTORY, https://www.justice.gov/open/file/1305831/dl.

[17] *E.g.*, Inseyets, CELLEBRITE, https://cellebrite.com/en/cellebrite-inseyets/; Magnet Axiom, MAGNET FORENSICS, https://www.magnetforensics.com/products/magnet-axiom/. There is a broad range of possible additional applications of AI for digital and multimedia forensics. *See* Johannes Fähndrich et al., *Digital Forensics and Strong AI: A Structured Literature Review*, 46 FORENSIC SCI. INT'L: DIGITAL INVESTIGATIONS (2023), https://doi.org/10.1016/j.fsidi.2023.301617.

[18] *E.g.*, Video Authentication Software, MEDEX FORENSICS, https://medexforensics.com/medex-platform/. The performance of AI-based tools for analyzing possible AI-generated or AI-modified content varies significantly, and the value of these tools depends on the nature of the evidence and the investigative context.

Future Uses of AI in Forensic Analysis

Recent and ongoing research has demonstrated a range of additional possible future applications of AI in forensic analysis.

a. Pattern and Trace Evidence

AI may be well suited for assisting experts in the comparison and categorization of some types of pattern and trace evidence.[19] Recent research has shown promise in applications of AI for assisting experts in analyzing toolmarks on bullets[20] and cartridges,[21] impressions of footwear outsoles,[22] fragments of glass,[23] traces of automotive paint,[24] and ignitable liquids,[25] among other types of evidence.

---

[19] Trace evidence involves material that has been transferred between objects, people, or the environment.

[20] *E.g.*, Eric Hare et al., *Algorithmic Approaches to Match Degraded Land Impressions*, 16 LAW, PROBABILITY & RISK 203-21 (2017), https://doi.org/10.1093/lpr/mgx018; Susan Vanderplas et al., *Comparison of Three Similarity Scores for Bullet LEA Matching*, 308 FORENSIC SCI. INT'L (2020), https://www.sciencedirect.com/science/article/pii/S0379073820300293; Pattranit Pisantanaroj et al., *Automated Firearm Classification From Bullet Markings Using Deep Learning*, 8 IEEE ACCESS 78236 (2020), https://ieeexplore.ieee.org/document/9076037.

[21] *E.g.*, Xiao Hui Tai & William F. Eddy, *A Fully Automatic Method for Comparing Cartridge Case Images*, 63 J. FORENSIC SCI. 440 (2018), https://onlinelibrary.wiley.com/doi/full/10.1111/1556-4029.13577; Joseph Roth et al., *Learning-based Ballistic Breech Face Impression Image Matching*, 2015 IEEE 7TH INTERNATIONAL CONFERENCE ON BIOMETRICS THEORY, APPLICATIONS AND SYSTEMS (2015), https://ieeexplore.ieee.org/document/7358774.

[22] *E.g.*, Soyoung Park & Alicia Carriquiry, *An Algorithm to Compare Two-dimensional Footwear Outsole Images Using Maximum Cliques and Speeded-up Robust Feature*, 13 STAT. ANALYSIS & DATA MINING: THE ASA DATA SCI. J. 188 (2020), https://onlinelibrary.wiley.com/doi/full/10.1002/sam.11449; Hana Lee et al., *An Automated Alignment Algorithm for Identification of the Source of Footwear Impressions with Common Class Characteristics*, 17 STAT. ANALYSIS & DATA MINING: THE ASA DATA SCI. J. (2024), https://dl.acm.org/doi/10.1002/sam.11659; Gautham Venkatasubramanian et al., *Quantitative Evaluation of Footwear Evidence: Initial Workflow for an End-to-end System*, 66 J. OF FORENSIC SCI. 2232 (2021), https://onlinelibrary.wiley.com/doi/full/10.1111/1556-4029.14802; Gautham Venkatasubramanian et al., *Comparing footwear impressions that are close non-matches using correlation-based approaches*, 66 J. FORENSIC SCI. 2232 (2021), https://onlinelibrary.wiley.com/doi/10.1111/1556-4029.14658; Moonsoo Jang & Soyoung Park, *A Finely Tuned Deep Transfer Learning Algorithm to Compare Outsole Images*, 16 STAT. ANALYSIS & DATA MINING: THE ASA DATA SCI. J. 511 (2023), https://onlinelibrary.wiley.com/doi/abs/10.1002/sam.11636; Zhijian Wen et al., *Shoeprint Image Retrieval and Crime Scene Shoeprint Image Linking by Using Convolutional Neural Network and Normalized Cross Correlation*, 63 SCI. & JUST. 439 (2023), https://doi.org/10.1016/j.scijus.2023.04.014; Samia Shafique et al., *CriSp: Leveraging Tread Depth Maps for Enhanced Crime-Scene Shoeprint Matching*, ARXIV (2024), https://arxiv.org/abs/2404.16972.

[23] *E.g.*, Omer Kaspi et al., *Toward Developing Techniques—Agnostic Machine Learning Classification Models for Forensically Relevant Glass Fragments*, 63 J. CHEMICAL INFO. & MODELING 87 (2022), https://pubs.acs.org/doi/full/10.1021/acs.jcim.2c01362; Grzegorz Zadora, *Glass Analysis for Forensic Purposes—A Comparison of Classification Methods*, 21 J. CHEMOMETRICS 174 (2007), https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/cem.1030.

[24] *E.g.*, Francis Kwofie et al., *Transmission Infrared Microscopy and Machine Learning Applied to the Forensic Examination of Original Automotive Paint*, 76 APPLIED SPECTROSCOPY 118 (2021), https://journals.sagepub.com/doi/10.1177/00037028211057574; George P. Affadu-Danful et al., *Raman Spectroscopy to Enhance Investigative Lead Information in Automotive Clearcoats*, 77 APPLIED SPECTROSCOPY 1064 (2023), https://journals.sagepub.com/doi/full/10.1177/00037028231186838.

[25] *E.g.*, Christian Bogdal et al., *Recognition of Gasoline in Fire Debris Using Machine Learning*, 331 FORENSIC SCI. INT'L (2022), https://doi.org/10.1016/j.forsciint.2021.111146; Michael E. Sigman et al., *Validation of Ground Truth Fire Debris Classification by Supervised Machine Learning*, 26 FORENSIC CHEMISTRY (2021), https://doi.org/10.1016/j.forc.2021.100358.

b. Drug Evidence

Analysis of seized drug samples is another promising area. Research has demonstrated that AI can assist with classifying fentanyl analogs and related compounds,[26] marijuana varieties,[27] and novel psychoactive substances.[28] These approaches generally combine established chemistry methods for identifying components of compounds, such as mass spectrometry, with machine learning methods to analyze the components and make categorizations.

c. Forensic Medicine, Pathology, and Anthropology

AI may assist with assessing injuries and injury mechanics. Analyzing a photograph of a bruise, for example, may enable estimating the date of the injury.[29]

AI may also be able to supplement expert analysis of human remains.[30] Recent publications show that it is feasible to estimate sex,[31] age,[32] and population affinity[33] from images and 2D and 3D computed tomography scans of skeletal and dental remains. AI may assist with identifying decedents, including through post-mortem iris recognition[34] and association of remains with

---

[26] *E.g.*, Travon Cooman et al., *Evaluation and Classification of Fentanyl-related Compounds Using EC-SERS and Machine Learning*, 68 J. FORENSIC SCI. 1520 (2023), https://onlinelibrary.wiley.com/doi/10.1111/1556-4029.15285; Phillip Koshute et al., *Machine Learning Model for Detecting Fentanyl Analogs from Mass Spectra*, 27 FORENSIC CHEMISTRY 100379 (2022), https://www.sciencedirect.com/science/article/abs/pii/S2468170921000758.

[27] *E.g.*, Austin McDaniel et al., *Toward the Identification of Marijuana Varieties by Headspace Chemical Forensics*, 11 FORENSIC CHEMISTRY 23-31 (2018), https://doi.org/10.1016/j.forc.2018.08.004.

[28] Swee Liang Wong et al., *Screening Unknown Novel Psychoactive Substances Using GC-MS based Machine Learning*, 34 FORENSIC CHEMISTRY 100499 (2023), https://www.sciencedirect.com/science/article/abs/pii/S2468170923000358?via%3Dihub.

[29] *E.g.*, Jhonatan Tirado & David Mauricio, *Bruise Dating Using Deep Learning*, 66 J. FORENSIC SCI. 336 (2020), https://pmc.ncbi.nlm.nih.gov/articles/PMC7821214/.

[30] *See* Laurent Tournois et al., *Artificial Intelligence in the Practice of Forensic Medicine: A Scoping Review*, 138 INT'L J. LEGAL MED. 1023 (2023), https://pmc.ncbi.nlm.nih.gov/articles/PMC11003914/; Nicola Galante et al., *Applications of Artificial Intelligence in Forensic Sciences: Current Potential Benefits, Limitations and Perspectives*, 137 INT'L J. LEGAL MED. 445 (2022), https://doi.org/10.1007/s00414-022-02928-5; Andrej Thurzo et al., *Use of Advanced Artificial Intelligence in Forensic Medicine, Forensic Anthropology and Clinical Anatomy*, 9 HEALTHCARE 1545 (2021), https://www.mdpi.com/2227-9032/9/11/1545; Micayla C. Spiros & Sherry Nakhaeizadeh, *We Think There's Been a Glitch: Artificial Intelligence and Machine Learning in Forensic Anthropology*, 7 FORENSIC ANTHROPOLOGY (2024), https://journals.upress.ufl.edu/fa/article/view/2827.

[31] Javier Venema et al., *Employing Deep Learning for Sex Estimation of Adult Individuals using 2D Images of the Humerus*, 35 NEURAL COMPUTING & APPLICATIONS 5987 (2023), https://link.springer.com/article/10.1007/s00521-022-07981-0; Tomoyuki Seo et al., *Sex Estimation Using Skull Silhouette Images from Postmortem Computed Tomography by Deep Learning*, 14 SCI. REPS. 22689 (2024), https://www.nature.com/articles/s41598-024-74703-y; Yongjie Cao et al., *Use of Deep Learning in Forensic Sex Estimation of Virtual Pelvic Models from the Han Population*, 7 FORENSIC SCIS. RES. 540 (2022), https://academic.oup.com/fsr/article/7/3/540/6987953.

[32] Juan Carlos Gámez-Granados et al., *Automating the Decision Making Process of Todd's Age Estimation Method from the Pubic Symphysis with Explainable Machine Learning*, 612 INFORMATION SCI. 514 (2022), https://www.sciencedirect.com/science/article/pii/S0020025522010301?via%3Dihub; NICHOLAS P. HERRMANN ET AL., INVESTIGATION OF SUBADULT DENTAL AGE-AT-DEATH ESTIMATION USING TRANSITION ANALYSIS AND MACHINE LEARNING METHODS (Off. Just. Progs. 2023), https://www.ojp.gov/pdffiles1/nij/grants/306558.pdf.

[33] David Navega et al., AncesTrees*: Ancestry Estimation with Randomized Decision Trees*, 129 INT'L J. LEGAL MED. 1145, 1145-1153 (2015),; G. Richard Scott et al., *rASUDAS: A New Web-Based Application for Estimating Ancestry from Tooth Morphology*, 1 FORENSIC ANTHROPOLOGY18, 18-31 (2018),.

[34] Aidan Boyd et al., *Post-Mortem Iris Recognition—A Survey and Assessment of the State of the Art*, 8 IEEE ACCESS 136570, 136570 (2020).

radiographs, dental records, photos, and 3D scans.[35] Where remains are incomplete, AI may help experts impute missing measurements.[36] Machine learning methods may also assist experts in categorizing and annotating images of remains,[37]and may be able to suggest certain causes of death, such as head trauma[38] or drowning.[39]

The ability of AI tools to effectively support forensic analysis in these ways will be contingent on the quality of photographs or imagery, among other factors.

d. Forensic Biology

AI has the promise of providing experts with new capabilities in probabilistic genetic analysis. These methods may improve the accuracy of genetic sequencing,[40] overcome gaps in degraded samples,[41] support inference of the number of contributors to a DNA mixture,[42] and enable predictions about physical appearance based on genetic information.[43]

---

[35] David C. Cornett et al., *Effects of Postmortem Decomposition on Face Recognition*, IEEE 10TH INT'L CONF. BIOMETRICS THEORY, APPLICATIONS, AND SYS. (2019), https://ieeexplore.ieee.org/document/9185971; A. Valsecchi et al., *Skeleton-ID: AI-driven Human Identification*, IEEE CONF. ARTIFICIAL INTEL. (2023), https://ieeexplore.ieee.org/abstract/document/10195123.

[36] Jinyong Pang & Xiaoming Liu, *Evaluation of Missing Data Imputation Methods for Human Osteometric Measurements*, 181 AM. J. BIOLOGICAL ANTHROPOLOGY 666 (2023), https://onlinelibrary.wiley.com/doi/10.1002/ajpa.24787.

[37] AUDRIS MOCKUS & DAWNIE WOLFE STEADMAN, ICPUTRD: IMAGE CLOUD PLATFORM FOR USE IN TAGGING AND RESEARCH ON DECOMPOSITION (Off. Just. Progs. 2020), https://www.ojp.gov/pdffiles1/nij/grants/255312.pdf; Sara Mousavi et al., *Machine-Assisted Annotation of Forensic Imagery*, 2019 IEEE INT'L CONF. ON IMAGE PROCESSING 1595 (2019), https://ieeexplore.ieee.org/document/8803068; Jack Garland et al., *Identifying Gross Post-mortem Organ Images Using a Pre-trained Convolutional Neural Network*, 66 J. FORENSIC SCI. 630 (2021), https://onlinelibrary.wiley.com/doi/10.1111/1556-4029.14608.

[38] Jack Garland et al., *Identifying Fatal Head Injuries on Postmortem Computed Tomography Using Convolutional Neural Network/Deep Learning: A Feasibility Study*, 65 J. FORENSIC SCIS., 2019 (2020), https://onlinelibrary.wiley.com/doi/full/10.1111/1556-4029.14502.

[39] Noriyasu Homma et al., *A Deep Learning Aided Drowning Diagnosis for Forensic Investigations Using Post-Mortem Lung CT Images*, 42ND ANNUAL INT'L CONF. OF THE IEEE ENG'G MED. & BIOLOGY SOC'Y (2020), https://ieeexplore.ieee.org/document/9175731.

[40] August E. Woerner et al., *Reducing Noise and Stutter in Short Tandem Repeat Loci with Unique Molecular Identifiers*, 51 FORENSIC SCI. INT'L: GENETICS 102459 (2021), https://www.fsigenetics.com/article/S1872-4973(20)30231-3/abstract; Michael A. Marciano et al., *A Hybrid Approach to Increase the Informedness of CE-based Data Using Locus-specific Thresholding and Machine Learning*, 35 FORENSIC SCI. INT'L: GENETICS 26 (2018), https://www.fsigenetics.com/article/S1872-4973(17)30313-7/abstract.

[41] Meng Huang et al., *A Machine Learning Approach for Missing Persons Cases with High Genotyping Errors*, 13 FRONTIERS GENETICS 1 (2022) https://www.frontiersin.org/journals/genetics/articles/10.3389/fgene.2022.971242.

[42] Michael A. Marciano & Jonathan D. Adelman, *PACE: Probabilistic Assessment for Contributor Estimation—A Machine Learning-based Assessment of the Number of Contributors in DNA Mixtures*, 27 FORENSIC SCI. INT'L: GENETICS 82 (2017), https://pubmed.ncbi.nlm.nih.gov/28040630/; Hamdah Alotaibi et al., *DNA Profiling: An Investigation of Six Machine Learning Algorithms for Estimating the Number of Contributors in DNA Mixtures*, 12 INT'L J. ADVANCED COMPUT. SCI. & APPLICATIONS 1 (2021), http://dx.doi.org/10.14569/IJACSA.2021.0121115.

[43] Maria-Alexandra Katsara et al., *Evaluation of Supervised Machine-learning Methods for Predicting Appearance Traits from DNA*, 53 FORENSIC SCI. INT'L: GENETICS 102507 (2021), https://doi.org/10.1016/j.fsigen.2021.102507.

Analysis of other types of biological evidence may also benefit from AI. Applying AI to microscopy images, for example, may assist experts with locating sperm cells in sexual assault evidence.[44]

Advances in forensic biology could be particularly valuable for investigations of violent crimes, where a perpetrator or victim may be more likely to leave biological evidence.

e. Forensic Toxicology

In toxicological analysis, AI may enable experts to identify unknown compounds.[45] Automated analysis may also be able to assess the toxicity of compounds, supporting treatment decisions made by medical professionals.[46]

f. Crime Scene Analysis

Complex crime scenes may involve thousands of photographs with nuanced attributes for forensic examiners to review and document. AI can support experts' work by automatically categorizing crime scene photographs based on visible items of interest, such as drugs and weapons.[47] AI may also be able to improve the quality of crime scene imagery for expert analysis, such as for scenes that are underwater.[48] In some instances, automated tools may be able to aid analysts in interpreting evidence within crime scene photographs, such as categorizing the potential mechanism that caused a blood spatter pattern.[49]

Challenges for AI in Forensic Analysis

The preceding chapter on Identification and Surveillance and the later chapter on Risk Assessment describe several challenges that are equally applicable to uses of AI in forensic analysis. In short, high-quality and representative data, rigorous and independent testing for performance and biases, ongoing monitoring, and established policies and oversight are all critical

---

[44] Raffael Golomingi et al., *Sperm Hunting on Optical Microscope Slides for Forensic Analysis with Deep Convolutional Networks – A Feasibility Study*, 56 FORENSIC SCI. INT'L: GENETICS 102602 (2022), https://www.sciencedirect.com/science/article/pii/S1872497321001393.

[45] Toshal D. Wankhade et al., *Artificial Intelligence in Forensic Medicine and Toxicology: The Future of Forensic Medicine*, 14 CUREUS (2022), https://doi.org/10.7759/cureus.28376; Zhoumeng Lin & Wei-Chun Chou, *Machine Learning and Artificial Intelligence in Toxicological Sciences*, 189 TOXICOL. SCI. 7 (2022), https://doi.org/10.1093/toxsci/kfac075.

[46] *See, e.g.*, Thi Tuyet Van Tran et al., *Artificial Intelligence in Drug Toxicity Prediction: Recent Advances, Challenges, and Future Perspectives*, 63 J. CHEM. INFO. & MODELING 2628 (2023), https://doi.org/10.1021/acs.jcim.3c00200.

[47] Joshua Abraham et al., *Automatically Classifying Crime Scene Images Using Machine Learning Methodologies*, 39 FORENSIC SCI. INT'L: DIGITAL INVESTIGATION 301273 (2021), https://doi.org/10.1016/j.fsidi.2021.301273; Amaljith Sreekumar et al., *Weapons and Related Object Classification in Digital Forensic Using Machine Learning*, 14TH INT'L CONF. ON COMPUTING COMMC'NS & NETWORKING TECHS. 1 (2023), https://ieeexplore.ieee.org/abstract/document/10307988.

[48] Rosella Paba et al., *Optimizing Underwater Visual Records for Crime Scene Investigations in Water with Clear to Reduced Visibility*, 6 FORENSIC SCI. INT'L: SYNERGY 100329 (2023), https://doi.org/10.1016/j.fsisyn.2023.100329.

[49] Yu Liu et al., *Automatic Classification of Bloodstain Patterns Caused by Gunshot and Blunt Impact at Various Distances*, 65 J. FORENSIC SCIS. 729 (2020), https://doi.org/10.1111/1556-4029.14262.

for uses of AI in criminal justice, including in support of forensic science. This section notes several challenges that could be particularly acute for the use of AI in forensic analysis.

a. Datasets

A large volume of high-quality data is essential for developing and evaluating AI uses in forensic analysis. If there is insufficient data, or if data has errors or gaps, that can undermine the potential value of AI for forensic science and create risks of harm from mistaken conclusions. Data that is properly distributed across a range of demographics and scenarios is also critical. This helps prevent bias in AI uses, leading to more consistent and fair forensic analysis.

Forensic analysis often depends on specialized data, such as data collected with dedicated equipment or from samples that are not readily available.[50] Preparing forensic datasets that are adequate for AI can also be expensive and labor intensive. Coordination by forensic experts worldwide to provide highly accurate and representative data may be essential for some uses of AI in forensic science.[51]

Respecting privacy is another important consideration for forensic science datasets. The data can be personal and sensitive, necessitating appropriate safeguards.

b. Validation

As noted earlier, analysis methods used in forensic science should be carefully studied for validity in principle, validity as applied in particular cases, and ongoing validity. Validation is essential for understanding the reliability and limitations of evidence analysis. Experts should be able to effectively characterize the accuracy and errors of AI used in forensic science, including the possibility of demographic disparities. When a forensic method involves AI, performance testing, bias testing, and continuous monitoring are all important steps for validation. Review of the implementation of an AI system, including the source code, can also be appropriate in some circumstances. At the federal level, OMB Memorandum M-24-10 requires testing and monitoring for certain use cases of AI used in support of forensic analysis. OMB Memorandum M-24-18 further requires use of independent testing data to the extent practicable, prohibits contractual restrictions on agency disclosure of testing methods and results, and encourages consideration of open-source AI implementations.

As an example, the path to widespread adoption of probabilistic genotyping highlights the importance and value of rigorous validation for new uses of AI in forensic science. When

[50] Because forensic analysis can involve specialized data, practitioners should be especially attentive to the risks of overfitting, where a model may learn correlations in training data that do not generalize to data in real-world uses. Overfitting can degrade performance and lead to unexpected model behaviors. At the federal level, OMB Memorandum M-24-10 specifically cautions agencies to consider the risks of overfitting.

[51] *See* Toshal D. Wankhade et al., *Artificial Intelligence in Forensic Medicine and Toxicology: The Future of Forensic Medicine*, 14 CUREUS 1, 4 (2022), https://doi.org/10.7759/cureus.28376. This type of data may be less necessary for certain types of analysis for digital and multimedia evidence, where the possible range of formats is determined by readily available software.

probabilistic genotyping began regularly appearing in forensic DNA analysis in the mid-2010s,[52] the President's Council of Advisors on Science and Technology (PCAST) expressed optimism that these methods could improve on prior approaches to analyzing samples with complex DNA mixtures, while also emphasizing the need for further validation.[53] PCAST concluded that studies had demonstrated validity in principle, but only for some mixtures of genetic material, and that additional steps were necessary to demonstrate validity as applied by laboratories. PCAST also expressed concern that validation studies had predominantly been carried out in collaboration with the vendors of probabilistic genotyping software. Scholars criticized probabilistic genotyping software as opaque and at risk of implementation errors, and they called attention to restrictive licensing agreements and trade secret protections that could inhibit independent validation.[54]

The forensic science community responded with a concerted effort to further substantiate probabilistic genotyping methods through independent and collaborative research. Laboratories worldwide coordinated in studies using the data available to each, reinforcing that particular methods could, in particular circumstances, have validity as they are deployed.[55] Vendors of probabilistic genotyping software improved access to their tools and source code for review by opposing legal teams.[56] While there is ongoing debate about the extent of validation for probabilistic genotyping—a 2021 draft comprehensive report by NIST concluded that gaps remain,[57] and a 2024 workshop by the National Academies reflected ongoing stakeholder concerns[58]—these steps have been positive and important.

---

[52] *See* SCIENTIFIC WORKING GROUP ON DNA ANALYSIS METHODS, GUIDELINES FOR THE VALIDATION OF PROBABILISTIC GENOTYPING SYSTEMS (2015), https://www.swgdam.org/_files/ugd/4344b0_22776006b67c4a32a5ffc04fe3b56515.pdf.

[53] PCAST Report, *supra* note 2.

[54] *E.g.*, Rebecca Wexler, *Life, Liberty, and Trade Secrets: Intellectual Property in the Criminal Justice System*, 70 STAN. L. REV. 1343, 1343 (2018), https://review.law.stanford.edu/wp-content/uploads/sites/3/2018/06/70-Stan.-L.-Rev.-1343.pdf; Andrea Roth, *Machine Testimony*, 126 YALE L.J. 1972, 1972 (2017), https://www.yalelawjournal.org/pdf/RothFinal_c4o97on1.pdf. *See also* Rediet Abebe et al., *Adversarial Scrutiny of Evidentiary Statistical Software*, 2022 ACM CONF. ON FAIRNESS, ACCOUNTABILITY, & TRANSPARENCY 1733 (2022), https://doi.org/10.1145/3531146.3533228.

[55] Safia Boodoosingh et al., *An Inter-laboratory Comparison of Probabilistic Genotyping Parameters and Evaluation of Performance on DNA Mixtures from Different Laboratories*, 71 FORENSIC SCI. INT'L: GENETICS 103046 (2024), https://doi.org/10.1016/j.fsigen.2024.103046; John M. Butler et al., *NIST Interlaboratory Studies Involving DNA Mixtures (MIX05 and MIX13): Variation Observed and Lessons Learned*, 37 FORENSIC SCI. INT'L: GENETICS 81 (2018), https://doi.org/10.1016/j.fsigen.2018.07.024; John M. Butler et al., *DNA Mixture Interpretation: A NIST Scientific Foundation Review*, NAT'L INST. STANDARDS & TECH. (2021), https://nvlpubs.nist.gov/nistpubs/ir/2021/NIST.IR.8351-draft.pdf (surveying validation studies of probabilistic genotyping software). *See* Jo-Anne Bright et al., *Internal Validation of STRmix™ – A Multi Laboratory Response to PCAST*, 34 FORENSIC SCI. INT'L: GENETICS 11 (2024), https://doi.org/10.1016/j.fsigen.2018.01.003.

[56] *E.g.*, *TrueAllele® Source Code Access*, CYBERGENETICS, https://www.cybgen.com/support/code-access/; *Access to STRmix™ Software By Defence Legal Teams*, STRMIX, https://www.strmix.com/assets/STRmix/STRmix-PDFs/Access-to-STRmix-Software-by-Defence-Legal-teams-March-2022-v2.pdf.

[57] John M. Butler et al., *DNA Mixture Interpretation: A NIST Scientific Foundation Review*, NAT'L INST. STANDARDS & TECH. (2021), https://nvlpubs.nist.gov/nistpubs/ir/2021/NIST.IR.8351-draft.pdf.

[58] NAT'L ACADS. SCIS., ENG'G & MED., *Law Enforcement Use of Probabilistic Genotyping, Forensic DNA Phenotyping, and Forensic Investigative Genetic Genealogy Technologies: Proceedings of a Workshop* (2024), https://doi.org/10.17226/27887.

These developments with probabilistic genotyping highlight the value of proactively carrying out large-scale interlaboratory validation for new uses of AI in forensic analysis. They also suggest that forms of access to AI implementation, such as models or source code, may be appropriate for validation.

c. Explainability

As noted earlier, it is important for forensics experts to be able to explain the analytical methods they apply and how they obtain results. The types of AI models that are used in forensic science today are generally interpretable, such that an expert could describe how inputs combine to arrive at an output.[59] It is foreseeable, though, that forensic science may begin to involve AI models that can be more difficult to understand, such as deep neural networks. These models could possibly be less persuasive in court and could undermine public confidence in forensic analysis. When considering models that are not interpretable, forensic practitioners should explore the feasibility of using explainability methods that can provide some understanding of model behavior.[60] Practitioners should also carefully consider possible tradeoffs between interpretability and accuracy in developing AI models for forensic science.

d. Human Oversight of AI in Forensic Analysis

In forensic science, it is common to have a second person review a practitioner's work, sometimes referred to as a "technical review." Similarly, when using AI, it is important to maintain human oversight of analysis and results to ensure that the AI was consistently applied and identify possible irregularities. Human involvement is also important because, if forensic analysis involving AI will be the basis for evidence in court, a human expert must explain the AI use and interpret the results.[61]

Including a human review element in the forensic analysis process comes with risks, however. Human review can introduce human biases, such as confirmation bias with respect to a subject or automation bias with respect to the reliability of analysis.[62] Training for forensic

---

[59] *See* Brandon L. Garrett & Cynthia Rudin, *Interpretable Algorithmic Forensics*, 120 PROCS. NAT'L ACAD. SCIS. (Oct. 2023), https://doi.org/10.1073/pnas.2301842120.

[60] *See* Louise Kelly et al., *Explainable Artificial Intelligence for Digital Forensics: Opportunities, Challenges and a Drug Testing Case Study*, DIGITAL FORENSIC SCIENCE (B. Suresh Kumar Shetty & Pavanchand Shetty eds., 2020), https://www.intechopen.com/chapters/73078; Stuart W. Hall & Amin Sakzad, *Explainable Artificial Intelligence for Digital Forensics*, 4 WIREs FORENSIC SCI. (2021), https://wires.onlinelibrary.wiley.com/doi/abs/10.1002/wfs2.1434; Marthe S. Veldhuis et al., *Explainable Artificial Intelligence in Forensics: Realistic Explanations for Number of Contributor Predictions of DNA Profiles*, 56 FORENSIC SCI. INT'L: GENETICS 102632 (2022), https://doi.org/10.1016/j.fsigen.2021.102632. *See generally* Alejandro Barredo Arrieta et al., *Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges Toward Responsible AI*, 58 INFO. FUSION 82 (2020), https://doi.org/10.1016/j.inffus.2019.12.012.

[61] *See* Andrea Roth, *Machine Testimony*, 126 YALE L.J. 1972 (2017), https://www.yalelawjournal.org/pdf/RothFinal_c4o97on1.pdf.

[62] *See* S. M. Kassin et al., *The Forensic Confirmation Bias: Problems, Perspectives, and Proposed Solutions*, 2 J. APPLIED RSCH. MEMORY & COGNITION 42 (2013), https://doi.org/10.1016/j.jarmac.2013.01.001. In forensic analysis of digital and multimedia evidence, for example, examiners often rely on tools that automate elements of analysis. The automation may not be robust against changes in the format of evidence, which could result from routine software updates. Automation bias could cause an examiner to miss relevant information if a tool is unsuccessful at

practitioners who use AI should address the risks of biases affecting analysis. Forensic analysis procedures should also minimize these risks, such as by minimizing the availability of irrelevant information to forensic examiners.[63]

Recommendations

a. Policies for Use of AI in Forensic Analysis

Forensic science service providers (FSSPs), such as law enforcement laboratories, should have clear and documented policies regarding the use of AI in forensic analysis. These policies should address the types of AI that may be used, the circumstances in which AI may be used, governance requirements, and limitations. FSSPs should consider using the NIST AI Risk Management Framework, NIST's Trustworthy and Responsible Artificial Intelligence Resource Center, OMB Memoranda M-24-10 and M-24-18, and other appropriate AI guidance from federal agencies to develop guidance and establish governance programs.

Consistent with accepted standards for forensic analysis, human review and interpretation of AI outputs should remain standard procedure in forensic science applications.

The output of an AI system should not be the sole basis for conclusions in forensic analysis. A qualified examiner should interpret the output and apply their expert judgment to form conclusions.

FSSPs should consider potential risks from using AI in forensic analysis and should design and implement processes to mitigate those risks prior to using AI. AI impact assessments may be a valuable framework for these considerations.

Policies should establish rigorous validation requirements to ensure that AI uses are reliable, both as they are developed and as they are deployed by the FSSP. Appropriate validation will often involve pre-deployment testing for performance and demographic biases, using data and contexts that are representative of real-world use, as well as post-deployment monitoring.

FSSP policies should address AI interpretability and explainability and should set a general preference for interpretable models when they can meet operational needs.

Regular audits of AI use can ensure that examiners are following required procedures.

b. Procurement of AI Capabilities for Forensic Analysis

FSSPs should only procure tools that have a demonstrated acceptable level of accuracy. FSSPs should verify that the data used to build an AI model is high quality and representative of the FSSP's intended real-world use.

extracting or parsing evidence. A possible mitigation for this risk would be to ensure that, if a tool encounters an error, it is clearly conveyed to examiners.

[63] Appropriate steps to minimize risks of human biases for forensic analysis involving AI may be similar or identical to appropriate steps for minimizing bias risks for analysis that does not involve AI.

FSSPs should require information from vendors about intended uses, training data and methods, validation, potential limitations, and potential biases of products. Experts should carefully review vendor disclosures and relevant additional information, such as peer-reviewed publications and reviews by other laboratories, for alignment with forensic analysis objectives.

Partnerships with independent AI researchers may be beneficial for FSSPs in validating AI products. Where possible, FSSPs should avoid restrictive licensing agreements and other possible barriers to collaboration with third-party experts.

FSSPs should ensure that biases in AI systems and uses, including for sex, race, color, disability, and age, have been adequately evaluated and mitigated.

c. Datasets

When possible, FSSPs should choose AI capabilities that have been trained on large, high-quality, and representative datasets. Producing these training datasets may require cooperation among practitioners and forensic experts worldwide, accounting for differences across jurisdictions that could affect performance or introduce biases. If an AI use depends on data that can differ in relevant ways across jurisdictions, FSSPs should consider appropriately supplementing training data with datasets from their jurisdictions.

When evaluating an AI system, FSSPs should also evaluate the training data where possible to ensure that the data is accurate, complete, and representative of the intended deployment context.

FSSPs should ensure that the data used to validate an AI system is separate from the data used to build the system. Reuse of training data in testing can lead to misunderstanding of the system's performance and biases.

d. Training and Education

Forensic practitioners who use or interact with AI should be appropriately trained about the AI, including its design, intended use, performance, biases, and limitations. Practitioners should also receive appropriate training about applicable policies and required procedures, as well as how to mitigate the possible human biases associated with use of the AI.

FSSPs can benefit each other by sharing information about their experiences with AI and the policies and procedures that they have implemented.

Forensic leaders should stay current on emerging AI tools and monitor advancements in forensic applications to take advantage of the technology and make changes to existing use of AI tools as necessary.

IV. Predictive Policing

Introduction

Predictive policing is the use of quantitative analytical methods to identify times, places, and individuals likely to be associated with criminal activity.[1] Predictive policing models do not predict that specific crimes will occur, but rather estimate a general likelihood of crime. While law enforcement agencies use predictive policing tools for purposes like informing the allocation of officers, existing tools neither recommend nor evaluate responses to their output. Those responses could extend beyond traditional policing methods and determine the efficacy of strategies incorporating predictive policing.

Law enforcement agencies have used predictive policing tools since at least the 1990s.[2] Researchers studied "hot spots" policing with the Minneapolis Police Department and found crime reductions in patrolled areas,[3] leading to further studies looking at approaches identifying crime "hot spots" to reduce drug and violent crime.[4] The Department of Justice's National Institute of Justice (NIJ) helped bring analysis tools to a wider range of U.S. law enforcement agencies by funding the development of CrimeStat software beginning in 1997.[5] By 2008, when approximately 90% of law enforcement agencies were already using some form of "hot spots" policing,[6] the NIJ joined with the Department of Justice's Bureau of Justice Assistance to fund research into new predictive models that would turn "hot spots" mapping into a forward-looking tool for crime forecasting.[7]

---

[1] WALTER L. PERRY ET AL., RAND CORP., PREDICTIVE POLICING: THE ROLE OF CRIME FORECASTING IN LAW ENFORCEMENT OPERATIONS 1–2 (2013), https://www.rand.org/content/dam/rand/pubs/research_reports/RR200/RR233/RAND_RR233.pdf. The term "predictive policing" is something of a misnomer because it is not a policing strategy, like proactive policing, problem-oriented policing, or crime prevention through environmental design. Instead, the outputs of predictive policing inform the development and implementation of policing or other strategies to reduce crime. However, this chapter uses the term "predictive policing" given its widespread acceptance. The chapter also references predictive policing algorithms, models, and tools to emphasize that predictive analytics are separate from the responses, which may include various forms of policing as well as community approaches to reducing crime.

[2] Lawrence W. Sherman & David Weisburd, *General Deterrent Effects of Police Patrol in Crime "Hot Spots": A Randomized, Controlled Trial*, 12 JUST. Q. 625, 643–46 (1995), http://dx.doi.org/10.1080/07418829500096221.

[3] *Id.*

[4] *See, e.g.*, Anthony A. Braga et al., *The Effects of Hot Spots Policing on Crime: An Updated Systematic Review and Meta-Analysis*, 31 JUST. Q. 633, 634 (2014).

[5] NED LEVINE, THE DEVELOPMENT OF A SPATIAL ANALYSIS TOOLKIT FOR USE IN A METROPOLITAN CRIME INCIDENT GEOGRAPHIC INFORMATION SYSTEM (1999) (Final Rep. to the Nat'l Inst. Just., Award No. 97-IJ-CX-0040), https://www.ojp.gov/pdffiles1/nij/grants/179282.pdf.

[6] *See* POLICE EXEC. RSCH. F., VIOLENT CRIME IN AMERICA: WHAT WE KNOW ABOUT HOT SPOTS ENFORCEMENT, 3 (2008), https://www.policeforum.org/assets/docs/Critical_Issues_Series/violent%20crime%20in%20america%20-%20what%20we%20know%20about%20hot%20spots%20enforcement%202008.pdf ("nearly 9 out of 10 agencies use hot spots enforcement efforts directed either at larger hotspots areas like neighborhoods, smaller hot spot places like intersections, or both").

[7] Joel Hunt, *From Crime Mapping to Crime Forecasting: The Evolution of Place-Based Policing*, NIJ J., no. 281, Nov. 2019, https://www.ojp.gov/pdffiles1/nij/252036.pdf.

While even early versions of predictive policing fall near this report's broad definition of artificial intelligence, predictive policing today uses more advanced statistical and machine learning methods and incorporates greater volumes and new types of data. Predictive policing tools can be grouped into two categories. "Location-based" or "place-based" tools attempt to forecast locations where crime is likely to cluster.[8] Location-based tools also try to identify the types of crimes likely to occur and when they are likely to occur. "Person-based" predictive policing attempts to identify individuals who are more likely to commit crimes or be victims of crime in the future, and it has many similarities to risk assessment (*see* chapter V).[9] The literature on predictive policing focuses overwhelmingly on "street crimes," and this chapter follows that research.

Predictive policing tools can help law enforcement agencies, other public services, and community-based organizations promote public safety, efficiency, and transparency by informing decisions on how to allocate limited resources. Nevertheless, these tools do not currently prescribe the actions that should be taken at the places or with the people identified—actions such as additional enforcement, place-based problem-solving, diversion programs, job training, education, or environmental design interventions—or predict the effects of those actions.

Law enforcement use of predictive policing also raises significant risks, including the potential to create or entrench disparities. The data underlying predictive policing models may have significant gaps and errors, and it may reflect historical and human biases. Use of predictive policing models based on that data may also result in unintended, unjust outcomes, such as over- or under-policing of certain individuals and communities.

This chapter begins with descriptions of uses and risks of both place-based and person-based predictive policing tools, and it concludes with recommendations for law enforcement agencies to safely and effectively deploy these tools in a way that enhances efficiency and accuracy while protecting civil liberties, civil rights, and privacy.

Uses of Predictive Policing

a. Place-based

Place-based or "hot spots" policing involves "focusing limited resources on a small number of high-activity crime places."[10] Also called "spatial models," promising analyses from Minneapolis and elsewhere suggested opportunities for these methods to assist in using policing resources more effectively and spurred extensive adoption by the 2000s.[11] During the 2010s, place-based predictive policing systems drew criticism about their accuracy, limitations, and reliance on historical crime.[12] As discussed below, predictive policing tools often rely heavily on historic crime

[8] *See* PERRY ET AL., *supra* note 1, at 8–9 (describing a similar taxonomy of predictive policing methods).
[9] *See id.*
[10] Anthony A. Braga, *Effects of Hot Spots Policing on Crime*, CAMPBELL SYSTEMATIC REVS. 1, 4 (2007).
[11] SHERMAN & WEISBURD, *supra* note 2, at 643–46; POLICE EXEC. RSCH. F., *supra* note 6.
[12] *See, e.g.*, Lyria Bennett Moses & Janet Chan, *Algorithmic Prediction in Policing: Assumptions, Evaluation, and Accountability*, 28 POLICING & SOC'Y 806, 809–13 (2018)(examining underlying assumptions in predictive policing

data, which can be problematic because the data may entrench historically discriminatory policing patterns. Criticisms also arose surrounding the immediate value and actionability of some systems' output for police departments.[13]

In 2016, the NIJ conducted the Real-Time Crime Forecasting Challenge to test the effectiveness and efficiency of predictive spatial models on crime data from Portland, Oregon.[14] Although results from the Challenge should not be generalized beyond the context of that single city, comparing the performance of a wide range of algorithms there indicated that both simple and sophisticated spatial models can offer similar predictive accuracy.[15] This finding supports the principle that for producing better outcomes, the specific tool or algorithm used for prediction may matter less than the selection and implementation of responses to high-risk areas.[16] For example, community engagement in the development of responses (e.g., additional police presence versus social services) could determine whether those responses are effective.[17]

Since the 2010s, place-based predictive policing strategies have continued to evolve, as agencies incorporate evidence from past iterations and work with communities to respond with a wide range of interventions. Agencies are developing tools in-house, customizing them to fit into their workflows, and collaborating with other public services and community-based organizations.[18] Strategies that currently make use of predictive policing include proactive policing and "hot spot" policing.[19] These strategies include a range of specific approaches to prevent crime and focus limited resources geographically.[20]

---

methods); DAVID ROBINSON & LOGAN KOEPKE, UPTURN, STUCK IN A PATTERN: EARLY EVIDENCE ON "PREDICTIVE POLICING" AND CIVIL RIGHTS 3–5 (2016), https://www.upturn.org/static/reports/2016/stuck-in-a-pattern/files/Upturn_-_Stuck_In_a_Pattern_v.1.01.pdf (discussing general limitations of predictive policing).

[13] *See* Aaron Sankin & Surya Mattu, *Predictive Policing Software Terrible at Predicting Crimes*, THE MARKUP (Oct. 2, 2023, 10:00 AM), https://themarkup.org/prediction-bias/2023/10/02/predictive-policing-software-terrible-at-predicting-crimes (finding low success rate of a predictive policing tool).

[14] *Real-Time Crime Forecasting Challenge*, NAT'L INST. JUST. (July 13, 2016), https://nij.ojp.gov/funding/real-time-crime-forecasting-challenge (archived content).

[15] *See* YongJei Lee et al., *A Theory-Driven Algorithm for Real-Time Crime Hot Spot Forecasting*, 23 POLICE Q. 174, 194–96 (2020) (analyzing data from Portland and Cincinnati and arguing that authors' simple model in Microsoft Excel "has demonstrated similar levels of efficiency and accuracy, with lower economic or fiscal investments and greater transparency than other more expensive, commercial models").

[16] *See, e.g.*, NAT'L ACAD. OF SCIS., ENG'G, & MED., LAW ENFORCEMENT USES OF PREDICTIVE POLICING APPROACHES (Nov. 2024) ("NAS Proceeding 2024"), https://nap.nationalacademies.org/catalog/28037/law-enforcement-use-of-person-based-predictive-policing-approaches-proceedings.

[17] NAS PROCEEDING 2024, *supra* note 16.

[18] *See* Tim Lau, *Predictive Policing Explained*, BRENNAN CTR FOR JUST. (Apr. 1, 2020), https://www.brennancenter.org/our-work/research-reports/predictive-policing-explained ("[T]he NYPD developed its own in-house predictive policing algorithms and started to use them in 2013").

[19] *See generally* Albert Meijer & Martijn Wessels, *Predictive Policing: Review of Benefits and Drawbacks*, 42 INT'L J. PUB. ADMIN. 1031, 1033–34 (2019) (discussing predictive policing and conventional policing methods).

[20] *See Practice Profile: Hot Spots Policing*, NAT'L INST. JUST. https://crimesolutions.ojp.gov/ratedpractices/hot-spots-policing#1-0 ("Hot spots policing strategies focus on small geographic areas or places, usually in urban settings, where crime is concentrated. Through hot spots policing strategies, law enforcement agencies can focus limited resources in areas where crime is most likely to occur.")

Recent place-based predictive policing programs include Place-Based Investigations of Violent Offender Territories (PIVOT) in Cincinnati[21] and Data-Informed Community Engagement (DICE) in cities such as Dallas, Kansas City, Newark, New Orleans, and St. Louis.[22] PIVOT uses "hot spot" mapping combined with problem-oriented policing strategies to identify and mitigate factors that facilitate violence,[23] while DICE pairs risk terrain modeling with community engagement to address crime through place-based interventions.[24]

b. Person-based

A person-based approach attempts to identify those most at risk for committing future crimes or being a victim of crime, either by using factors associated with individuals known to law enforcement to generate risk scores, or by identifying connections between individuals who may be linked to past crimes. This approach typically generates a list of the highest-risk individuals in the jurisdiction as a whole or within a given geographic area, such as a patrol zone. With respect to violent crime, research consistently shows that it is disproportionately concentrated among small numbers of individuals, groups, and locations at the highest risk for violence.[25]

In the mid-1990s, the NIJ funded an early study of one form of predictive policing, focused deterrence, in which law enforcement efforts prioritized individuals disproportionately responsible for crime.[26] In the past decade, predictive policing programs like Chicago's Strategic Subjects List built on earlier person-focused deterrence approaches and other initiatives, such as gang databases.[27] That generation of policing programs established the use of algorithms to analyze data on individuals' criminal histories, social networks, and other risk factors in an attempt to identify those most likely to be involved in violent crime as perpetrators or victims. By the end of the

[21] *Place-Based Investigations of Violent Offender Territories (PIVOT)*, CITY OF CINCINNATI, https://www.cincinnati-oh.gov/police/community-involvement/pivot/.
[22] Learning Community, PUB. SAFETY COLLABORATIVE COUNCIL, https://www.diceforpublicsafety.org/learning-community.html.
[23] *See* Tamara D. Madensen et al., *Place-Based Investigations to Disrupt Crime Place Networks*, THE POLICE CHIEF MAG., Apr. 2017, at 14–15, https://www.policechiefmagazine.org/wp-content/uploads/PoliceChief_April2017_F_WEB.pdf (describing development and implementation of PIVOT).
[24] Sarah Minster, *The Data-Informed Community Engagement (DICE) Approach to Public Safety Turns Analytics into Action*, NAT'L LEAGUE OF CITIES (June 18, 2024), https://www.nlc.org/article/2024/06/18/the-data-informed-community-engagement-dice-approach-to-public-safety-turns-analytics-into-action.
[25] *Violent Crime Reduction Roadmap: Working Together to Build Safer Communities*, Action 2, BUREAU JUST. ASSISTANCE, https://bja.ojp.gov/violent-crime-reduction-roadmap/action-2#0-3 ("Research consistently shows that violence disproportionately concentrates among small number of individuals, groups and locations at the highest risk for violence.")
[26] Braga et al. 2001, *Problem-Oriented Policing, Deterrence, and Youth Violence: An Evaluation of Boston's Operation Ceasefire*, J. RSCH. CRIME & DELINQ. 195, 198 (2001), https://journals.sagepub.com/doi/abs/10.1177/0022427801038003001 (discussing the Boston Gun Project and Operation Ceasefire).
[27] *See generally* Jessica Saunders et al., *Predictions Put into Practice: A Quasi-Experimental Evaluation of Chicago's Predictive Policing Pilot*, 12 J. EXPERIMENTAL CRIMINOLOGY 347 (2016), https://link.springer.com/article/10.1007/s11292-016-9272-0.

2010s, however, many of the programs from this era had halted following concerns regarding efficacy, bias, and civil rights.[28]

Person-based strategies predicting risk for victimization rather than criminal offending are also being used in cases of child abuse[29] and gender-based violence.[30] These tools raise some similar concerns to offender-centered predictive policing tools.[31]

Risks of Predictive Policing

a. Data and Model Output Quality

Although traditional law enforcement interventions have relied on officers' knowledge, judgments, and personal experience in their communities, predictive modeling relies on collected quantitative data as input. Predictive models depend on historic crime data, such as calls for service, crime incidents, and arrests.[32] As with all data, these datasets are imperfect.[33] They may underrepresent underreported crimes, such as intimate partner violence and sexual assault.[34] They may also reflect past policing patterns, which can affect the certainty of the model's output and disproportionately affect vulnerable communities.[35]

Predictive models may also underrepresent the actual incidence of crime in locations where individuals are less likely to report crimes to police.[36] Some scholars have pointed to historical

---

[28] *See, e.g.*, Annie Sweeney & Jeremy Gorner, *For Years Chicago Police Rated the Risk of Tens of Thousands Being Caught Up in Violence. That Controversial Effort Has Quietly Been Ended.*, CHI. TRIB. (Jan. 25, 2020, 2:55 AM), https://www.chicagotribune.com/2020/01/24/for-years-chicago-police-rated-the-risk-of-tens-of-thousands-being-caught-up-in-violence-that-controversial-effort-has-quietly-been-ended/ (discussing a findings in a report issued by the City Inspector General's Office analyzing the use of Chicago's Strategic Subjects List).

[29] *See, e.g.*, ALLEGHENY CNTY. DEP'T HUM. SERVS., SUMMARIZING RECENT RESEARCH ON PREDICTIVE RISK MODELS IN CHILD WELFARE 1 (2024), https://www.alleghenycountyanalytics.us/wp-content/uploads/2024/05/24-ACDHS-04-Predictive-Risk-Algorithms.pdf (discussing the Allegheny Family Screening Tool, "an algorithm designed to assist child welfare call screening caseworkers in their assessment of general protective service referrals regarding potential child maltreatment.").

[30] *See, e.g.*, ETICAS FOUND., THE EXTERNAL AUDIT OF THE VIOGÉN SYSTEM (2022), https://eticasfoundation.org/wp-content/uploads/2024/07/ETICAS-FND-The-External-Audit-of-the-VioGen-System-1.pdf. (examining VioGén, a gender-violence risk assessment tool used by the Spanish Ministry of Interior).

[31] ALLEGHENY CNTY. DEP'T HUM. SERVS., *supra* note 29.

[32] *See, e.g.*, Jeffrey Brantingham et al., *Does Predictive Policing Lead to Biased Arrests? Results From a Randomized Controlled Trial*, STATISTICS AND PUBLIC POLICY 5(1), 1–6. https://doi.org/10.1080/2330443X.2018.1438940 (evaluating bias of predictive algorithms used for police patrol using arrest data).

[33] *Id.* at 5 *(*discussing limitations of arrest data).

[34] *See* U.S. DEP'T JUST, FRAMEWORK FOR PROSECUTORS TO STRENGTHEN OUR NATIONAL RESPONSE TO SEXUAL ASSAULT AND DOMESTIC VIOLENCE INVOLVING ADULT VICTIMS 2 n. 8 (May 2024), https://www.justice.gov/ovw/media/1352371/dl?inline ("Sexual assaults and domestic violence are, in large part, underreported, under-investigated and under-prosecuted").

[35] Some observers have pointed to racial disparities in policing decisions in such instances. *See, e.g.*, AM. C. L. UNION N.J., SELECTIVE POLICING RACIALLY DISPARATE ENFORCEMENT OF LOW-LEVEL OFFENSES IN NEW JERSEY 8 (Dec. 2015), https://www.aclu-nj.org/sites/default/files/field_documents/2015_12_21_aclunj_select_enf.pdf.

[36] The Bureau of Justice Statistics reported that in 2022, only 41.5% of violent crimes and 31.8% of property crimes were reported to police. ALEXANDRA THOMPSON & SUSANNAH N. TAPP, BUREAU OF JUST. STATS., U.S. DEP'T JUST., NCJ 307089, CRIMINAL VICTIMIZATION, 2022 6 tbl.4 (2023), https://bjs.ojp.gov/document/cv22.pdf.

public health statistics as another way to estimate the underlying incidence of crime[37] just like the Drug Enforcement Administration (DEA) has used public-facing CDC drug overdose data to evaluate and to more effectively address the drug threat in communities across the United States.[38] Looking further afield, other crime prediction models have incorporated a wide variety of data sourced from outside law enforcement and unconnected to past criminal conduct, including "elevation; zoning and water area coverage; density of hospitals, fire departments, transportation points such as bus stops and subway entrances, and schools," as well as variables such as the day of the week and the weather.[39] Many datasets, including data on crime incidents, may also contain some uncertainty because data features such as the time of a crime or the location of a crime may have a large range or be inaccurately reported.

Police officers typically generate data each time they conduct a stop, write a report, or do anything else that leaves a data trail. Some police actions are subject to discretion, and the discretion of individual officers thus plays a role in shaping the input data.[40] As the next section discusses, the use of AI in predictive policing runs the risk of introducing or exacerbating disparities in this input data.

b. Civil Rights

Perhaps the most frequent criticism of predictive policing is that it has the potential to reproduce or even amplify biases embedded in historical crime data. The risk is that historical crime data will train models to predict crime in ways that magnify biases.[41] For example, place-based models can drive increased police presence in an area with greater historical crime data. The increased presence can lead to more law enforcement activity in that area, which can result in even more officers assigned to the area. Ultimately, the models become "increasingly confident that the locations most likely to experience further criminal activity are exactly the locations they had previously believed to be high in crime."[42] The result may be disparate outcomes, by race or other

---

[37] *See* Kristian Lum & William Isaac, *To Predict and Serve?*, SIGNIFICANCE 15, 16–17 (Oct. 2016) https://rss.onlinelibrary.wiley.com/doi/pdf/10.1111/j.1740-9713.2016.00960.x (comparing police records of drug crimes with public health survey of illegal drug use, which the authors argue is a better "ground truth" for incidence of drug crimes).

[38] *See*, *e.g.*, Press Release, Drug Enforcement Admin., DEA Launches New Initiative to Combat Drug-Related Violence and Overdoses in Communities Across America (Feb. 7, 2022), https://www.dea.gov/press-releases/2022/02/07/dea-launches-new-initiative-combat-drug-related-violence-and-overdoses-0 ("DEA initiated a data-driven approach using national crime statistics and CDC data to identify hot spots of drug-related violence and overdose deaths across the country, in order to devote its law enforcement resources where they will have the most impact: the communities where criminal drug networks are causing most harm."); DEA Administrator Anne Milgram Remarks as Delivered Press Conference (Dec. 16, 2021) https://www.dea.gov/sites/default/files/2021-12/DEA%20Administrator%20Anne%20Milgram%20Remarks%20as%20Delivered-Dec%2016%202021.pdf (discussing CDC data on drug overdose deaths and criminal drug network activities).

[39] Jerry H. Ratcliffe et al., *The Philadelphia Predictive Policing Experiment*, 17 J. EXPERIMENTAL CRIMINOLOGY 15, 21 (2021) (describing possible inputs for predictive policing model formerly known as HunchLab); *see also* Resourcerouter Frequently Asked Questions, SoundThinking, https://www.soundthinking.com/faqs/resourcerouter-faqs/ ("We supplement data modeling with non-crime data . . . . Typical examples include seasonality, time of month, day of the week, time of day, holidays, upcoming events, weather, and locations of liquor establishments.").

[40] *See* Lum & Isaac, *supra* note 38, at 16 (describing "feedback loop" in predictive model trained on police records that overrepresent drug crimes in heavily patrolled areas).

[41] *Id.* at 19.

[42] *Id.* at 16.

demographic characteristics, for residents of that area even where rates of crime may be comparable to other neighborhoods.

Person-based models may also be trained on data reflecting underlying disparities,[43] which may adversely impact individuals who receive increased attention and law enforcement interactions. In one person-based program, "officers are instructed to focus their attention on the highest point-value individuals … [who] are subjected to heightened surveillance, and, therefore, are more likely to be stopped, thus further increasing their point value."[44]

This situation can create feedback loops in which "targeting of certain areas or certain races creates the impression of higher crime rates in those areas, which then justifies continued police presence there."[45] Beyond the impact on residents and individuals, feedback loops can erode public trust in police by increasing police presence, and thus enforcement activities, in communities that are already distrustful of law enforcement.

c. Privacy

Individuals that a predictive policing tool assesses to be more likely to engage in or become a victim of crime may be subject to additional law enforcement scrutiny, raising concerns about privacy, surveillance, and harassment. Even given stringent criteria for inclusion of an individual's data in a predictive policing system, the absence of oversight and monitoring for adherence to these criteria can mean that some individuals experience law enforcement surveillance or interactions without substantiated links to criminal activity. Precedent exists for oversight concerns with criminal justice databases.[46]

Developers of predictive policing models may also choose to augment police data by drawing from additional datasets.[47] However, merging other datasets may also raise privacy concerns, such as revealing sensitive details of individuals or communities.[48]

---

[43] *See* Andrew Guthrie Ferguson, *Policing Predictive Policing*, 94 WASH. U. L. REV. 1109, 1148–49 (2017) ("The result has been to justify disproportionate minority contacts and the collection of minority names in databases. These actions then feed a confirmation feedback loop that equates those currently in the system with those who need to be policed by the system. Essentially, high-crime areas or high-value suspects might only be considered 'high' because police already have data about those areas or people.").

[44] Sarah Brayne & Angèle Christin, *Technologies of Crime Prediction: The Reception of Algorithms in Policing and Criminal Courts*, Social Problems 1, 13 (2020) https://doi.org/10.1093/socpro/spaa004; *see also* Sarah Brayne, *Big Data Surveillance: The Case of Policing*, 977, 987 (2017), https://doi.org/10.1177/0003122417725865 ("An individual having a high point value is predictive of future police contact and that police contact further increases the individual's point value.").

[45] Ferguson, *supra* note 43 at 1153.

[46] *See, e.g.*, CAL. STATE AUDITOR, THE CALGANG CRIMINAL INTELLIGENCE SYSTEM 1, 31–32 (2016), https://www.auditor.ca.gov/pdfs/reports/2015-130.pdf (finding California gang database's "oversight structure does not ensure that user agencies collect and maintain criminal intelligence in a manner that does not invade individuals' privacy rights," and that some agencies lacked adequate support for including individuals in database).

[47] *See* Ratcliffe et al., *supra* note 3938.

[48] *E.g.* Arvind Narayanan & Vitaly Shmatikov, Robust De-anonymization of Large Sparse Datasets, https://ieeexplore.ieee.org/document/4531148 (connecting Netflix and IMDB data to learn potentially sensitive video-watching habits of individuals).

d. Interpreting and Explaining

As models become more complex, they gain the ability to represent more intricate relationships and sophisticated patterns in the data.[49] This increased flexibility, however, often comes with increased difficulty for humans to interpret and explain the internal workings of these complex models.[50]

In predictive policing, more complex models are likely to be marginally more accurate than simpler ones,[51] especially when predicting rarer crimes. However, the diminishing marginal improvements that come with more sophisticated algorithms may not justify their use over simpler ones, given their higher cost, greater complexity, and decreased interpretability. A less complex predictive policing tool that is more interpretable may better align with law enforcement and community needs and values than a highly complex "black box" tool. An interpretable tool can allow for greater accountability and oversight by providing concrete explanations for a model's predictions and may even increase transparency into the factors that influenced decisions.

Recommendations

Thoughtful deployment of predictive policing tools can not only mitigate risks but can also ensure that AI use for predictive policing offers accuracy and efficiency benefits for law enforcement. Entities that have deployed or are considering deploying predictive policing tools should consider the following recommendations.

a. Assess Goals of the Predictive Policing Tool with the Community

When considering a predictive policing tool, a critical question is the particular community's goals that have driven consideration of the tool. After identifying objectives, a community should outline how it will measure the success of the tool or alternatives in addressing those goals. These initial steps should engage public agencies, relevant community-based organizations, and other stakeholders in the community, such that agencies can determine how to align their use of tools to community needs and expectations, address concerns and potential risks, and establish consensus for the implementation process. Failure to effectively engage with impacted communities can undermine public trust. Law enforcement agencies weighing the adoption of predictive policing tools should therefore engage with community members and their representatives (e.g., county commissioners or city council members) regarding goals and measures of success.

b. Assess the Need for a Predictive Policing Tool and Possible Alternatives

Law enforcement agencies should assess the likelihood a predictive policing tool will address the community's goals, the shortcomings the tool might have, alternatives to the tool

[49] *See, e.g.*, PERRY ET AL., *supra* note 1, at 36 (comparing "simple methods," such as regression analyses, to "black box models," which can model "extremely complicated relationships").
[50] Whereas simple models "are usually directly interpretable by a person," more complex "black box models" are not. PERRY ET AL., *supra* note 1, at 36.
[51] *See, e.g.*, PERRY ET AL., *supra* note 1, at xix ("Although there is usually a correlation between the complexity of a model and its predictive power, increases in predictive power have tended to show diminishing returns.").

(including non-AI alternatives), and ways the tool might address or exacerbate concerns. Along with these considerations, law enforcement agencies should evaluate whether they have the resources, personnel, and technical expertise necessary for the proper use and monitoring of these types of systems in the short and long term.

This assessment—and the full lifecycle of a predictive policing tool's deployment—should include community engagement. Law enforcement agencies should seek to educate community members and their representatives on how a considered predictive policing tool works, the benefits of the tool, the short- and long-term costs of the tool, the risks associated with the tool (e.g., civil rights concerns), and plans for monitoring and mitigating risks. When consulting with stakeholders, law enforcement agencies should provide the public meaningful opportunities for participation by using plain language, considering language access needs, and communicating with those most likely to be negatively impacted by these tools.

c. Assess Which Data to Use to Train Models and Ensure Data Is As Accurate As Possible

Law enforcement agencies should carefully consider which data to include in a predictive policing model. To mitigate the risk of feedback loops, agencies should keep the model inputs focused on types of crime relevant to their goals. Additionally, because models should be continually updated with new data, agencies should consider filtering out data that reflect actions taken by police in response to previous model outputs in order to avoid feedback loops.

Agencies should also ensure a central role for humans in choosing which data to use as inputs, deciding which metrics to use for evaluating accuracy, screening the model's outputs with a crime analyst's eye for contextual knowledge and historical factors, and determining which practices to implement in response to predictions. Although humans have the potential to introduce their own biases, they are also essential for ensuring that predictive policing tools are being used according to law and regulation, as well as relevant policies. By being transparent with the public about these choices, especially with respect to the selection of input data for predictive tools, agencies can ensure accountability and build public trust.

Predictive models may have a veneer of neutrality that can be reinforced by automation bias.[52] To counteract this, user interfaces for predictive policing systems should include constraints, cues, notifications, and embedded content that reinforce appropriate use of the model, especially with regard to critically evaluating model outputs. Training users to think critically about the limits to algorithmic objectivity and error rates must be a top priority in any law enforcement organization using predictive models.[53]

---

[52] "Automation bias refers to undue deference to automated systems by human actors that disregard contradictory information from other sources or do not (thoroughly) search for additional information." Saar Alon-Barkat & Madaline Busuioc, *Human–AI Interactions in Public Sector Decision Making: "Automation Bias" and "Selective Adherence" to Algorithmic Advice*, 33 J. PUB. ADMIN. RSCH. & THEORY 153, 155 (2023).

[53] *See* ALEXANDER BABUTA & MARION OSWALD, ROYAL UNITED SERVS. INST. FOR DEF. & SEC. STUD., DATA ANALYTICS AND ALGORITHMIC BIAS IN POLICING 15 (2019), https://static.rusi.org/20190916_data_analytics_and_algorithmic_bias_in_policing_web_0.pdf ("Adequate training

Although elimination of all errors from a large dataset may be infeasible, law enforcement agencies have a responsibility to ensure that the data they collect on both incidents and individuals are as accurate as possible. Routine checks of police and other criminal justice data for accuracy and ongoing revisions to improve data quality can mitigate the risks associated with data errors for both the police and community members. To the extent feasible, organizations using predictive policing tools should establish channels for affected community members to determine what data is stored about them, learn about decisions made based on that data, and petition for omission or corrections of errors.

d. Test, Measure, Validate, and Reevaluate, Including Independently and in a Real-world Context

For validation purposes, agencies should use data collected before and after deployment. Pre-deployment data are recommended because they isolate the system's ability to forecast without any interference from changes in response to the predictions. Collecting post-deployment data is recommended because they might help in measuring the efficiency or accuracy of the model over time when responses are changing and potentially detecting, deterring, or preventing crimes in the forecasted areas.[54]

After validating the model, agencies should also conduct operational and field testing to better understand the impact of the model combined with interventions. Even accurate and well-validated models could produce unintended consequences when implemented in practice, so agencies should have mechanisms in place to identify the broad impact of their prediction-based interventions—including the possibility of bias and discrimination—and make any appropriate changes based on continual evaluation.[55]

The impact of a particular predictive policing approach on crime rates is a commonly applied metric, but other factors can yield a more complete picture of the impact, including: community-defined measures of success; citizen commendations or complaints; use of excessive or unwarranted force; traffic stops and field interviews; citations, fines, and fees; victim satisfaction surveys; citizen surveys about factors such as fear of crime and satisfaction with and trust in police; number of tips received; rates of victim and witness cooperation; officer health, wellbeing, retention, and job satisfaction; and geographic residency of new applicants to the police force.[56] To identify disparate treatment or impacts, evaluations of models should include data

---

focused on cognitive bias and fair decision-making would appear essential to ensure officers are able to consistently achieve the correct balance.").

[54] "Deployment of prediction boxes to the field and delivery of policing dosage to those boxes is expected to suppress some fraction of crime in those locations. Predictive accuracy should therefore decline in response to directed patrol." G.O. Mohler et al., *Randomized Controlled Field Trials of Predictive Policing*, 110 J. AM. STAT. ASS'N 1399, 1404 (2015).

[55] *See, e.g.*, NAT'L INST. STANDARDS & TECH., ARTIFICIAL INTELLIGENCE RISK MANAGEMENT FRAMEWORK (AI RMF 1.0), at 33 tbl.4 (2023), https://doi.org/10.6028/NIST.AI.100-1.

[56] Research exploring the accuracy of place-based predictive policing algorithms commonly relies on the Prediction Accuracy Index (PAI) or the Prediction Efficiency Index* (PEI*). *See* Spencer Chainey et al., *The Utility of Hotspot Mapping for Predicting Spatial Patterns of Crime*, 21 SEC. J. 4 (2008) (proposing PAI); Veronica M. White et al., *A Discussion of Current Crime Forecasting Indices and an Improvement to the Prediction Efficiency Index for*

disaggregated by demographic factors such as race, national origin, age, gender, disability, and other characteristics.

Testing and validation are necessary regardless of whether an agency develops a predictive policing tool in-house or acquires the tool from a third party. Part of testing and evaluation should include completing an impact assessment for the predictive policing tool before putting it into use.

Agencies should adopt a mechanism for objective third-party auditing of models and source code to proactively address concerns about model accuracy, reliability, and potential for bias and discrimination. Agencies should routinely audit police databases to mitigate the costs of data errors for both users of the system and community members. Databases used in person-based predictive policing, for example, should be routinely audited to ensure that individuals meet the criteria for being included in the database.

Models should be reevaluated periodically based on the policing patterns and crime rates that result from their use, ensuring the models are fair; equitable, including minimizing feedback loops; and aligned with public safety priorities. If seeking to procure a predictive policing tool from a third party, agencies should require regular evaluation and auditing as part of the vendor contract.

Importantly, agencies should mitigate emerging risks to rights and safety, including by regularly updating the predictive policing tool to improve it and reduce its risks.

Metrics used to assess accuracy in person-based predictive policing have much in common with those used in risk assessment, because person-based tools estimate, among other things, the likelihood of a person being involved in a crime. See chapter V for a detailed discussion of risk assessment.

e. Broadly Consider Community Resources to Address Root Causes of Crime

How predictive policing tools are used matters.[57] In addition to (or in place of) a policing response, community-based interventions built around resources for public health and social services can help alleviate the burdens on officers and mitigate concerns about predictive models.[58]

f. Adopt Policies

Law enforcement agencies should adopt policies that address the use of the predictive policing system pre- and post-deployment. Such policies should govern which predictive policing systems are approved for use, as well as how predictive models are selected, trained, implemented,

---

Applications, 37 SEC. J. 47 (2024); *see also* Ned Levine, *The "Hottest" Part of a Hotspot: Comments on "The Utility of Hotspot Mapping for Predicting Spatial Patterns of Crime"*, 21 SEC. J. 295 (2008) (discussing Recapture Rate Index (RRI)); Veronica M. White & Joel Hunt, *Measuring How Relatively "Good" a Hot-spot Map Is: A Summary of Current Metrics*, 1 PROCEEDINGS OF THE IISE ANNUAL CONFERENCE & EXPO 2022, at 97 (2022), available at https://www.ojp.gov/pdffiles1/nij/305365.pdf (evaluating PAI, RRI, PEI, and PEI*).

[57] *Cf.* EXEC. OFF. OF THE PRESIDENT, BLUEPRINT FOR AN AI BILL OF RIGHTS 53–54 (Oct. 2022), https://www.whitehouse.gov/wp-content/uploads/2022/10/Blueprint-for-an-AI-Bill-of-Rights.pdf (describing "predictive policing" as an application of AI that may impact civil rights, civil liberties, or privacy).

[58] ANDREW GUTHRIE FERGUSON, THE RISE OF BIG DATA POLICING: SURVEILLANCE, RACE, AND THE FUTURE OF LAW ENFORCEMENT 173–76 (2017).

interpreted, and evaluated. Each of these policies should consider the tradeoffs between public safety, community trust, and the potential harms of algorithmic error, human bias, automation bias, and system inequalities that may result in misclassification. For further discussion of governance measures, see the Conclusion and Best Practices chapter.

g. Ensure Adequate Human Training and Assessment

Law enforcement organizations using predictive policing tools should require regular training on the operation and appropriate use of the specific tools. The training should be appropriate for the tools' diverse users, and it should include modules to ensure users think critically about the limits to algorithmic objectivity, error rates, ways in which these tools are embedded in existing inequitable systems, and strategies to prevent and mitigate human biases and systemic inequities.

Users of predictive policing models should receive training on how to interpret the model's outputs. The required training should also address how to determine what type of intervention, from police or other services, would be best to prevent or address the types of crime predicted in a given community, especially if predictions are driven by public health factors like substance abuse, homelessness, and mental health disorders.

Law enforcement personnel should not be authorized to use predictive policing systems if their training is not current.

V. Risk Assessment

Risk assessment instruments use quantitative analysis to estimate the likelihood that a certain outcome will occur in the criminal justice system, such as whether an individual will recidivate or fail to appear for trial.[1] These instruments are commonly based on statistical models.[2] Agencies and courts may use the model output to inform decisions about individuals, including whether to require detention before trial, what sentence to impose, and what interventions to attempt.

Risk assessment in criminal justice dates back nearly a century,[3] and these types of tools are currently widely used.[4] Based on the data available, it appears that every state implements some form of risk assessment,[5] with considerable local variation,[6] and most states have a state law or court rule that addresses risk assessment.[7] At the federal level, risk assessment instruments

[1] *See* P'SHIP ON AI, REPORT ON ALGORITHMIC RISK ASSESSMENT TOOLS IN THE U.S. CRIMINAL JUSTICE SYSTEM 7 (2019), https://partnershiponai.org/paper/report-on-machine-learning-in-risk-assessment-tools-in-the-u-s-criminal-justice-system/ (discussing definitions of risk assessment).

[2] As noted in the introduction, this report follows the definition of "artificial intelligence" from OMB Memoranda M-24-10 and M-24-18, which does not depend on "the type of model" and includes "simple" models, such as regression models that use conventional statistics.

[3] BERNARD E. HARCOURT, AGAINST PREDICTION: PROFILING, POLICING, AND PUNISHING IN AN ACTUARIAL AGE 47-76 (2013), https://academic.oup.com/chicago-scholarship-online/book/21282.

[4] *See Risk Assessment Landscape*, PUB. SAFETY RISK ASSESSMENT CLEARINGHOUSE, https://bja.ojp.gov/program/psrac/selection/risk-assessment-landscape (surveying uses of risk assessment nationwide); *Risk Assessment Tool Database*, BERKMAN KLEIN CTR., https://criminaljustice.tooltrack.org/ (compiling data from reports and jurisdictions about uses of risk assessment), *How Many Jurisdictions Use Each Tool?*, MAPPING PRETRIAL INJUSTICE, https://pretrialrisk.com/national-landscape/how-many-jurisdictions-use-each-tool/ (summarizing survey data about uses of risk assessment); *Stanford Pretrial Risk Assessment Tools Factsheet Project*, STANFORD L. SCH. POL'Y LAB, https://law.stanford.edu/pretrial-risk-assessment-tools-factsheet-project/ (collecting authoritative descriptions of common risk assessment instruments); NAT'L CTR. FOR STATE CTS., APPENDIX A: PROFILES OF ASSESSMENT INSTRUMENTS, https://www.ncsc.org/__data/assets/pdf_file/0014/27140/ran-appendix-a.pdf (same); *Pretrial Release: Risk Assessment Tools*, NAT'L CONF. OF STATE LEGISLATURES (June 30, 2022), https://www.ncsl.org/civil-and-criminal-justice/pretrial-release-risk-assessment-tools (surveying state laws and court rules that address the development and use of risk assessment instruments); *50-State Report on Public Safety Part 2, Strategy 2, Action Item 2*, THE COUNCIL OF STATE GOVERNMENTS, https://50statespublicsafety.us/part-2/strategy-2/action-item-2/ (providing results of a 50-state survey on use of risk assessment in probation and parole); Cathy Hu et al., *National Scan of Policy and Practice in Risk Assessment Policy Brief Number 2017-01*, THE RISK ASSESSMENT CLEARINGHOUSE (July 20, 2018, 3:13 PM), https://bja.ojp.gov/sites/g/files/xyckuh186/files/media/document/PB-Scan-of-Practice.pdf (survey of risk assessment practices, policies, procurement, and perceptions in 43 states).

[5] *50-State Report on Public Safety Part 2, Strategy 2, Action Item 2*, THE COUNCIL OF STATE GOVERNMENTS, https://50statespublicsafety.us/part-2/strategy-2/action-item-2/ (providing results of a 50-state survey on use of risk assessment in probation and parole).

[6] *Risk Assessment Landscape*, PUB. SAFETY RISK ASSESSMENT CLEARINGHOUSE, https://bja.ojp.gov/program/psrac/selection/risk-assessment-landscape (surveying uses of risk assessment nationwide).

[7] NAT'L CONF. OF STATE LEGISLATURES (June 30, 2022), https://www.ncsl.org/civil-and-criminal-justice/pretrial-release-risk-assessment-tools (surveying state laws and court rules that address the development and use of risk assessment instruments).

inform decision-making by courts, court services, and prisons. The Model Penal Code also encourages the use of risk assessment.[8]

This chapter begins with an overview describing how risk assessment is currently used in critical stages of the criminal justice system, as well as the typical design of these instruments. Next, the chapter describes the potential benefits of quantitative risk assessment in making criminal justice more effective, equitable, and efficient, followed by a discussion of risks including inaccuracy, biases, errors in data, inadequate validation, and insufficient consideration of alternatives. The chapter closes with a set of recommendations for the development and use of risk assessment instruments.

Uses of Risk Assessment

Risk assessment instruments inform decisions throughout the criminal justice system, including the following phases of prosecution and detention.[9] For each phase, this chapter discusses how risk assessment is used and provide examples of specific risk assessment tools currently in use. As discussed in greater detail below, risk assessment tools should be employed only after a thorough evaluation, including for predictive performance,[10] potential bias,[11] and suitability for the subject population, as well as the implementation of appropriate precautions.

a. Pretrial Release

Pretrial service agencies and prosecutors make recommendations to judges and judicial officers who ultimately make decisions about the status of defendants before trial.[12] A defendant may be released with no or minimal conditions, released subject to supervision and conditions, required to post bail to be released, or detained. Risk assessment can inform these decisions by assisting decision-makers in estimating the likelihood that a defendant will fail to appear in court, commit another offense before trial, or pose a risk to public safety. Examples used at the state and local levels include the Virginia Pretrial Risk Assessment Instrument (VPRAI) and VPRAI

[8] Am. Law Inst., MODEL PENAL CODE: SENTENCING § 6.03 reporter's note F (AM. L. INST., Proposed Final Draft 2017), available at https://robinainstitute.umn.edu/sites/robinainstitute.umn.edu/files/2022-02/mpcs_proposed_final_draft.pdf; *id.* at § 6.09, cmt. E.

[9] *See* Melissa Hamilton, *Risk Assessment Tools in the Criminal Legal System – Theory and Practice: A Resource Guide*, Nat'l Ass'n of Crim. Def. Laws. 16 fig. 1 (Nov. 2020), https://www.nacdl.org/getattachment/a92d7c30-32d4-4b49-9c57-6c14ed0b9894/riskassessmentreportnovember182020.pdf (noting a range of criminal justice system decision points where risk assessment instruments are used).

[10] As discussed further below, there are a range of possible predictive performance metrics for risk assessment models, including precision and the false positive rate.

[11] Similarly, as discussed further below, there are a range of possible bias metrics for risk assessment models, including calibration and predictive equality. *See generally* Alessandro Castelnovo et al., *A Clarification of the Nuances in the Fairness Metrics Landscape*, SCIENTIFIC REPORTS (2022); Simon Caton & Christian Haas, *Fairness in Machine Learning: A Survey*, ACM COMP. SURVEYS (2024).

[12] In the federal court system, federal laws establish a general presumption that a defendant should be released before trial unless the government establishes at a detention hearing that the defendant should be detained because they present a risk to the community or are unlikely to appear at subsequent proceedings. Many state court systems also have a presumption of release and allow for defendants to be released if they post cash bail.

Revised (VPRAI-R),[13] Public Safety Assessment (PSA),[14] Ohio Risk Assessment System Pretrial Assessment Tool (ORAS-PAT),[15] and Correctional Offender Management Profiling for Alternative Sanctions Pretrial Release Risk (COMPAS PRRS-I and -II).[16] In the federal system, U.S. Probation and Pretrial Services uses the Pretrial Risk Assessment (PTRA), an algorithmic tool developed by the Administrative Office of the U.S. Courts.[17]

b. Sentencing

Following a criminal conviction, a judge often has discretion to determine an appropriate sentence for the defendant, subject to applicable statutory ranges and sentencing guidelines.[18] Risk assessment tools can support these decision-makers in estimating the likelihood that a defendant will recidivate, which is typically a factor in sentencing. Widely used risk assessments in sentencing include the Level of Service Revised (LSI-R) and Level of Service Case Management Inventory (LS/CMI),[19] as well as the COMPAS General Recidivism Risk Scale (GRRS), Violent Recidivism Risk Scale (VRRS), and COMPAS-R Summative GRRS.[20] Some jurisdictions use

---

[13] *See* MARIE VANNOSTRAND & KENNETH J. ROSE, VA. DEP'T OF CRIM. JUST. SERVS., PRETRIAL RISK ASSESSMENT IN VIRGINIA, VIRGINIA PRETRIAL RISK ASSESSMENT INSTRUMENT (2009), https://www.dcjs.virginia.gov/sites/dcjs.virginia.gov/files/publications/corrections/virginia-pretrial-risk-assessment-report.pdf; VA. DEP'T OF CRIM. JUST. SERVS., VIRGINIA PRETRIAL RISK ASSESSMENT INSTRUMENT - (VPRAI): INSTRUCTION MANUAL – VERSION 4.5 (2021), https://www.dcjs.virginia.gov/sites/dcjs.virginia.gov/files/publications/corrections/virginia-pretrial-risk-assessment-instrument-vprai_2.pdf; KENNETH ROSE, STANFORD L. SCH. POL'Y LAB, RISK ASSESSMENT FACT SHEET: VIRGINIA PRETRIAL RISK ASSESSMENT INSTRUMENT (VPRAI) (2019), https://law.stanford.edu/wp-content/uploads/2019/06/VPRAI-Factsheet-FINAL-6-20.pdf.

[14] *See* ADVANCING PRETRIAL POL'Y & RSCH, *About the Public Safety Assessment*, https://advancingpretrial.org/psa/about/; KRISTIN BECHTEL (ARNOLD VENTURES), STANFORD L. SCH. POL'Y LAB, RISK ASSESSMENT FACT SHEET: PUBLIC SAFETY ASSESSMENT (PSA), https://law.stanford.edu/wp-content/uploads/2019/05/PSA-Sheet-CC-Final-5.10-CC-Upload.pdf.

[15] Edward J. Latessa et al, *The Creation and Validation of the Ohio Risk Assessment System (ORAS)*, 74 Fed. Prob. J. 16 (2010), https://www.uscourts.gov/sites/default/files/74_1_2_0.pdf.

[16] EQUIVANT, PRACTITIONER'S GUIDE TO COMPAS CORE (Apr. 2019) https://web.archive.org/web/20190520172536/http://www.equivant.com/wp-content/uploads/Practitioners-Guide-to-COMPAS-Core-040419.pdf.

[17] *Pretrial Risk Assessment*, ADMIN. OFF. OF THE U.S. CTS., https://www.uscourts.gov/services-forms/probation-and-pretrial-services/supervision/pretrial-risk-assessment.

[18] In several jurisdictions, a jury may determine the appropriate sentence. Research on how risk assessment interacts with human decision makers, discussed below, has focused on judges rather than juries.

[19] Christopher T. Lowenkamp & Kristin Bechtel, *The Predictive Validity of the LSI-R on a Sample of Offenders Drawn from the Records of the Iowa Department of Corrections Data Management System*, 71 FED. PROB. J. 25 (2007) https://www.uscourts.gov/sites/default/files/71_3_4_0.pdf; JAMES AUSTIN ET AL., INST. ON CRIME, JUST. & CORR. AT GEO. WASH. UNIV., *Reliability and Validity Study of the LSI–R Risk Assessment Instrument, Final Report submitted to the Pennsylvania Board of Probation and Parole* (Jan. 9, 2003) https://www.ojp.gov/ncjrs/virtual-library/abstracts/reliability-and-validity-study-lsi-r-risk-assessment-instrument.

[20] Eugenie Jackson & Christina Mendoza, Equivant/Northpointe, *Setting the Record Straight: What the COMPAS Core Risk and Need Assessment Is and Is Not,* HARVARD DATA SCI. REV., Mar. 31, 2020 https://hdsr.mitpress.mit.edu/pub/hzwo7ax4/release/7; EQUIVANT, PRACTITIONER'S GUIDE TO COMPAS CORE (Apr. 2019) https://web.archive.org/web/20190520172536/http://www.equivant.com/wp-content/uploads/Practitioners-Guide-to-COMPAS-Core-040419.pdf; Antonio Cordella & Francesco Gualdi, *Algorithmic Formalization: Impacts on Administrative Processes*, PUB. ADMIN. (Aug. 27, 2024), https://onlinelibrary.wiley.com/doi/10.1111/padm.13030.

specialized instruments to evaluate the risk of recidivism in defendants convicted of sexual offenses.[21]

c. Prison Classification

If a defendant is in custody or has a custodial sentence imposed, a court, jail, prison, or supporting agency must make decisions about detention. These decisions can include the type of facility, housing unit assignment, placement in a general or special population, and availability of programs, services, and work. Risk assessment instruments can inform these decisions by calling attention to predictors of potential violence and other misconduct while in custody.[22] The tools used for prison classification are predominantly designed to estimate the risk of recidivism after release,[23] though some have been modified or evaluated for predicting prison misconduct.[24] Tools

---

[21] R. Karl Hanson et al., *Assessing the Risk and Needs of Supervised Sexual Offenders: A Prospective Study Using STABLE-2007, Static-99R, and Static-2002R*, 42 CRIM. JUST. & BEHAV., Dec. 2015, at 1205, 1205, https://journals.sagepub.com/doi/full/10.1177/0093854815602094 (evaluating risk assessment tools applied to individuals convicted of sexual offenses).

[22] *See generally* NAT'L INST. OF CORR., U.S. DEP'T JUST., OBJECTIVE PRISON CLASSIFICATION: A GUIDE FOR CORRECTIONAL AGENCIES (2nd ed. 2021), https://nicic.gov/resources/nic-library/all-library-items/objective-prison-classification-guide-correctional-agencies; Richard A. Berk et al., *A Randomized Experiment Testing Inmate Classification Systems*, 2 CRIM. & PUB. POL'Y 215 (2003) (describing a randomized study to evaluate a prison classification system); Patricia L. Hardyman et al., *Internal Prison Classification Systems: Case Studies in Their Development and Implementation*, NAT'L INST. OF CORR. (Jan. 2002), https://nicic.gov/resources/nic-library/all-library-items/internal-prison-classification-systems-case-studies-their (describing the prison classifications systems in several states); Joe Russo, Michael J.D. Vermeer, Dulani Woods & Brian A. Jackson, *Risk and Needs Assessments in Prisons: Identifying High-Priority Needs for Using Evidence-Based Practices*, RAND (Sept. 9, 2020), https://www.rand.org/pubs/research_reports/RRA108-5.html; Daryl G. Kroner & Jeremy F. Mills, *The Accuracy of Five Risk Appraisal Instruments in Predicting Institutional Misconduct and New Convictions*, 28 CRIM. JUST. & BEHAV., Aug. 2001, at 471, 471, https://journals.sagepub.com/doi/10.1177/009385480102800405; Thomas R. Kane, *The Validity of Prison Classification: An Introduction to Practical Considerations and Research Issues*, 32 CRIME & DELINQ., July 1986, at 367, 367, https://journals.sagepub.com/doi/abs/10.1177/0011128786032003008 (describing methods for validating prison classification systems).

[23]James M. Byrne & Amy Dezember, *The Research Director Perspective on the Design, Implementation, and Impact of Risk Assessment and Offender Classification Systems in USA Prisons: A National Survey*, *in* HANDBOOK ON RISK AND NEED ASSESSMENT 48, 53 (Faye Taxman ed., 2016), https://www.taylorfrancis.com/chapters/edit/10.4324/9781315682327-10/research-director-perspective-design-implementation-impact-risk-assessment-offender-classification-systems-usa-prisons-national-survey-james-byrneand-amy-dezember?context=ubx&refId=9495cf84-b71e-455f-9547-8e80237d7fa7 (reporting results of a multistate survey on prison classification systems and finding that most responding states used a risk assessment tool designed to predict recidivism).

[24] Edward Latessa et al., *Creation and Validation of the Ohio Risk Assessment System: Final Report*, UNIV. OF CINCINNATI CTR. FOR CRIM. JUST. RSCH. (July 2009), https://www.uc.edu/content/dam/uc/ccjr/docs/reports/project_reports/ORAS_Final_Report.pdf (describing the ORAS Prison Intake Tool, which is designed for prison intake use and validated as an estimate of post-release recidivism); Joshua S. Long, Appropriate Classification of Prisoners: Balancing Prison Safety with the Least Restrictive Placements of Ohio Inmates (Jun. 22, 2020) (Ph.D. dissertation, University of Cincinnati), https://cech.uc.edu/content/dam/refresh/cech-62/school-of-criminal-justice/research/2020/Joshua%20S.%20Long%206-22-20.pdf; Matthew Makarios & Edward J. Latessa, *Developing a Risk and Needs Assessment Instrument for Prison Inmates: The Issue of Outcome*, 40 CRIM. JUST. & BEHAV., Dec. 2013, at 1449, 1449, https://journals.sagepub.com/doi/full/10.1177/0093854813496240 (comparing risk assessment tools in the prison classification context, one designed to estimate misconduct in custody and the other designed to estimate post-release recidivism, and concluding that are significant differences and a "one size fits all" approach to risk assessment may undermine validity).

that have been specifically designed and validated for predicting prison misconduct are less common.[25] In the federal system, the Federal Bureau of Prisons currently uses the Prisoner Assessment Tool Targeting Estimated Risk and Needs (PATTERN)[26] to track dynamic changes in risk and the Standardized Prisoner Assessment for Reduction in Criminality (SPARC-13)[27] to identify programmatic and treatment needs of inmates.

d. Probation, parole, and supervision

Risk assessment tools are commonly used to inform decisions about the appropriate level of supervision for convicted persons who are not in custody, conditions for release and reentry plans from custody, and eligibility for earned release from custody.[28] Assessments that attempt to calculate likelihood of recidivism are common in these settings, too, such as LSI-R and COMPAS GRRS. In the federal system, the Federal Bureau of Prisons uses PATTERN to inform eligibility for earned time.[29] The Administrative Office of the U.S. Courts also developed the Post Conviction Risk Assessment (PCRA), which aims to predict general and violent recidivism and is used by federal probation officers.[30]

Risk Assessment Design

There are, broadly, two types of risk assessment instruments: actuarial models, which quantitatively combine factors to estimate the likelihood that a risk will occur, and structured

[25] *E.g.*, Grant Duwe, *The Development and Validation of a Classification System Predicting Severe and Frequent Prison Misconduct*, 100 THE PRISON J., Mar. 2020, at 173 https://journals.sagepub.com/doi/10.1177/0032885519894587 (describing the design and validation of a risk assessment tool for estimating the likelihood of misconduct in custody); Mark D. Cunningham et al., *An Actuarial Model for Assessment of Prison Violence Risk Among Maximum Security Inmates*, 12 ASSESSMENT, Mar. 2005, at 40 https://journals.sagepub.com/doi/abs/10.1177/1073191104272815 (similar).

[26] NAT'L INST. OF JUST., 2023 REVIEW AND REVALIDATION OF THE FIRST STEP ACT RISK ASSESSMENT TOOL (Aug. 2024), https://www.ojp.gov/pdffiles1/nij/309264.pdf; Zachary Hamilton et al., *Tailoring to a Mandate: The Development and Validation of the Prisoner Assessment Tool Targeting Estimated Risk and Needs (PATTERN)*, 39 JUST. Q., Apr. 2021, at 1129, 1129, https://www.tandfonline.com/doi/full/10.1080/07418825.2021.1906930.

[27] NAT'L INST. OF JUST., NCJ 309349, 2023 REVIEW AND VALIDATION OF THE FEDERAL BUREAU OF PRISONS NEEDS ASSESSMENT SYSTEM (Sept. 2024), https://www.ojp.gov/pdffiles1/nij/309349.pdf.

[28] *See, e.g.*, John Monahan & Jennifer L. Skeem, *Risk Assessment in Criminal Sentencing*, 12 ANN. REV. CRIM. PSYCH. 489, 494, 496 (2016) (describing risk assessments to shorten a sentence on the "back end"); Sheldon X. Zhang et al., *An Analysis of Prisoner Reentry and Parole Risk Using COMPAS and Traditional Criminal History Measures*, 6 Crime & Delinquency 167 (2014) (describing use of COMPAS in parole supervision).

[29] *See* U.S. GOV'T ACCOUNTABILITY OFF., PUB. NO. GAO-23-105139, BUREAU OF PRISONS SHOULD IMPROVE EFFORTS TO IMPLEMENT ITS RISK AND NEEDS ASSESSMENT SYSTEM, https://www.gao.gov/assets/gao-23-105139.pdf (describing how PATTERN is used for earned time eligibility).

[30] *Post Conviction Risk Assessment*, ADMIN. OFF. OF THE U.S. CTS., https://www.uscourts.gov/services-forms/probation-and-pretrial-services/supervision/post-conviction-risk-assessment; *see also* Seena Fazel et al., *The Predictive Performance of Criminal Risk Assessment Tools Used at Sentencing: Systematic Review of Validation Studies*, 81 J. CRIM. JUST., July–Aug. 2022 (analyzing sentencing and post-conviction tools).

judgments,[31] which provide frameworks for applying experience and intuition.[32] The simplest actuarial designs compute a sum of factors, without any weighting, to generate an overall risk score. More complex approaches use conventional statistical methods, typically linear or logistic regression. Recent models use advanced statistics and machine learning methods, such as boosted regression[33] or gradient-boosted decision trees.[34] Typically, the quantitative output of a model is converted into a category, such as low, moderate, or high risk, by applying predefined thresholds.[35]

The features that risk assessments take into consideration also vary significantly. Some have over 100 distinct inputs, while others consider fewer than 10.[36] The general long-term trend is toward fewer features that have a more readily understandable relationship with outcomes in the criminal justice system.[37] Early risk assessments tended to emphasize subjective characterizations of defendants, derived from professional interviews.[38] Some early risk assessment tools also relied on impermissible factors, such as race.[39]

Modern risk assessment instruments tend to employ more objective factors that are more closely related to criminal justice outcomes. The factors may be fixed (e.g., criminal history) or changeable over time (e.g., time since last infraction or participation in drug treatment).[40] Some recent risk assessment tools also account for stages of the criminal justice system and periodic

[31] Structured judgments are not a focus of this report, since they fall outside the definition of artificial intelligence in OMB Memoranda M-24-10 and M-24-18. This chapter discusses them below as an important comparison for actuarial models, which can be AI within the OMB definition.
[32] Sarah L. Desmarais & Jay P. Singh, *Risk Assessment Instruments Validated and Implemented in Correctional Settings in the United States*, THE COUNCIL OF STATE GOV'TS JUST. CTR. (Mar. 27, 2013), https://csgjusticecenter.org/wp-content/uploads/2020/02/Risk-Assessment-Instruments-Validated-and-Implemented-in-Correctional-Settings-in-the-United-States.pdf (surveying the design and validation of risk assessment tools); Jennifer L. Skeem & John Monahan, *Current Directions in Violence Risk Assessment*, VA. PUB. L. & LEGAL THEORY RSCH. PAPER NO. 2011-13 (Mar. 2011), https://gspp.berkeley.edu/assets/uploads/research/pdf/03-2011_Current_Directions_in_Violence_Risk_Assessment.pdf .
[33] Zachary Hamilton et al., *Tailoring to a Mandate: The Development and Validation of the Prisoner Assessment Tool Targeting Estimated Risk and Needs (PATTERN)*, 39 JUST. Q., Apr. 2021, at 1129, 1129, https://www.tandfonline.com/doi/full/10.1080/07418825.2021.1906930.
[34] Jon Kleinberg et al., *Human Decisions and Machine Predictions*, 133 Q.J. ECON., Aug. 2017, at 237, 237, https://academic.oup.com/qje/article/133/1/237/4095198.
[35] The thresholds are typically set when developing, validating, or revalidating a risk assessment tool. The thresholds and their quantitative and qualitative meanings vary by risk assessment tool.
[36] See the references accompanying the Uses of Risk Assessment section for factors that widely used risk assessment tools analyze.
[37] Bernard E. Harcourt, *Risk as a Proxy for Race: The Dangers of Risk Assessment*, 27 FED. SENT'G REP. 237 (2015), https://scholarship.law.columbia.edu/cgi/viewcontent.cgi?article=3568&context=faculty_scholarship (describing the long-term historical trend in risk assessment factors).
[38] James Bonta & D.A. Andrews, *Risk-Need-Responsivity Model for Offender Assessment and Rehabilitation*, PUBLIC SAFETY CANADA (Jan. 2007), https://www.publicsafety.gc.ca/cnt/rsrcs/pblctns/rsk-nd-rspnsvty/rsk-nd-rspnsvty-eng.pdf ; D.A. Andrews, James Bonta & J. Stephen Wormith, *The Recent Past and Near Future of Risk and/or Need Assessment*, 52 CRIME & DELINQ., Jan. 2006, at 7, 7, https://journals.sagepub.com/doi/10.1177/0011128705281756 ; Desmarais & Singh *supra* note 32.
[39] Early risk assessment instruments considered race as a factor. Modern models do not. *See* Bernard E. Harcourt, *Risk as a Proxy for Race: The Dangers of Risk Assessment*, 27 FED. SENT'G REP. 237 (2015), https://scholarship.law.columbia.edu/cgi/viewcontent.cgi?article=3568&context=faculty_scholarship.
[40] These different categories of factors are sometimes delineated as "static" and "dynamic."

reassessment and are designed to interact with treatment and supervision conditions that can affect risk calculations.[41]

Future risk assessment approaches could incorporate more frequent model validation and updates, as well as more frequent reassessments based on evolving factors.[42] These changes could improve predictive performance and reduce biases. For example, researchers are currently exploring how to evaluate changes in risk based on real-time location and provide prompt support to individuals based on their unique risks and needs.[43] These types of ongoing assessments may create privacy risks for individuals, and future approaches may have to further account for both maximizing predictive performance and respecting privacy.[44]

Potential Benefits for a More Effective, Equitable, and Efficient Criminal Justice System

The central promise of risk assessment is that empirical evaluation of the risk of future harmful behavior could be more accurate, transparent, and equitable than subjective human judgments alone.

More accurate assessments of risk have the potential for significant benefits. They could enable better alignment between justice system functions including rehabilitation, deterrence, and incapacitation, and the prevention of future offenses.[45] Better predictions could also more accurately identify people who are unlikely to reoffend, channeling them toward lesser pretrial restrictions, lesser sentences, and less restrictive conditions of release.[46]

Efficiency in resource allocation is another possible upside of more accurate estimates of risk likelihood.[47] Costlier aspects of criminal justice monitoring or detention could be better

[41] The 2022 annual review and validation of PATTERN, for example, found that individuals in federal prison could and often did see their risk level change from first to last assessment, independent of simply getting older (which is associated with lower risk). NAT'L INST. OF JUST., NCJ 305720, 2022 REVIEW AND REVALIDATION OF THE FIRST STEP ACT RISK ASSESSMENT TOOL 18, 35 (Mar. 2022), https://www.ojp.gov/pdffiles1/nij/305720.pdf.
[42] *See* D. Michael Applegarth et al., *Imperfect Tools: A Research Note on Developing, Applying, and Increasing Understanding of Criminal Justice Risk Assessments*, 34 CRIM. JUST. POL'Y REV. 319, 323 (2023), https://journals.sagepub.com/doi/epub/10.1177/08874034231180505 (noting support of dynamic factors by focus group of winners of the National Institute of Justice's Recidivism Forecasting Challenge).
[43] The National Institute of Justice recently funded research to develop a real-time, cellphone-based intelligent tracking system to monitor people who are on community supervision with the goal of flagging potentially risky behavior and providing support to avert such actions. *See* MARCUS ROGERS, NCJ 308693, AI ENABLED COMMUNITY SUPERVISION FOR CRIMINAL JUSTICE SERVICES 2–3 (2024) (Final Rep. to the Nat'l Inst. of Just., Award No. 2019-75-CX-K001), https://www.ojp.gov/pdffiles1/nij/grants/308693.pdf.
[44] *See id.* at 12–13, 36.
[45] *See* KLEINBERG ET AL., *supra* note 33 at 237; Erin Collins, *Punishing Risk*, 107 GEO. L.J. 57, 72-73 (2018), https://www.law.georgetown.edu/georgetown-law-journal/wp-content/uploads/sites/26/2018/12/Punishing-Risk-2.pdf; John Monahan & Jennifer L. Skeem, *Risk Assessment in Criminal Sentencing*, 12 ANN. REV. CLINICAL PSYCH., 489, 489, https://doi.org/10.1146/annurev-clinpsy-021815-092945.
[46] *See* Sarah L. Desmarais, John Monahan & James Austin, *The Empirical Case for Pretrial Risk Assessment Instruments*, CRIM. J. & BEHAV. (2021).
[47] Erin Collins, *Punishing Risk*, 107 GEO. L.J. 57, 76-77 (2018), https://www.law.georgetown.edu/georgetown-law-journal/wp-content/uploads/sites/26/2018/12/Punishing-Risk-2.pdf..

directed toward situations where there is greater predicted benefit.[48] For a given level of public resources, it may be possible to obtain a greater level of public safety.[49]

Risk assessment instruments also have the potential to increase transparency of human judgments.[50] A risk assessment tool could be made publicly accessible, along with its design documentation, validation studies, and guidance to practitioners.[51] The data used for risk assessment could provide the basis for review of decisions—and correction, if necessary—by affected individuals and their counsel. A subjective decision-maker, by contrast, might not (intentionally or otherwise) fully explain the information that they considered and how they arrived at a decision.

Equity is another important motivation for using risk assessment tools. Models can be designed and validated to minimize disparities in predictive performance across demographic groups.[52] Models can also be designed and validated to ensure that individuals with similar salient characteristics, such as the crime charged and criminal history, receive similar estimates of risk likelihood.[53]

When risk assessments are appropriately designed, validated, and used, actuarial models can be predictive of criminal justice outcomes and can outperform human judgments alone.[54]

---

[48] There are, to be sure, other factors to consider in allocating criminal justice resources. Crime prevention is an important goal, but not the only goal.

[49] KLEINBERG ET AL. *supra* note 33.

[50] Alex Chohlas-Wood, *Understanding Risk Assessment Instruments in Criminal Justice*, BROOKINGS (June 19, 2020), https://www.brookings.edu/articles/understanding-risk-assessment-instruments-in-criminal-justice/.

[51] The Public Safety Assessment (PSA), for example, is available to the public. So are the PSA's design documentation, validation studies, and guidance to practitioners.

[52] This conception of equity is, in AI research and practice, often referred to as "group fairness." See the design documentation and validation studies referenced in the first section of this chapter for examples of how risk assessment tool developers address equity considerations. As discussed further below, there are differing possible equity metrics, and risk assessment tools may be considered equitable by some metrics and not by others. The factors used for risk assessment can also be closely related to demographics (e.g., residential and income data can be related to race), adding further complexity to the challenge of measuring and mitigating biases.

[53] In AI research and practice, this conception of equity is often referred to as "individual fairness." As discussed further below, the use of "cut points" with risk assessment tools can introduce or exacerbate risks to individual fairness. *See* Jane R. Bambauer, Tal Zarsky & Jonathan Mayer, *When a Small Change Makes a Big Difference: Algorithmic Fairness Among Similar Individuals*, 55 U.C. DAVIS L. REV. 2337 (2022).

[54] SHAMENA ANWAR ET AL., RAND CORP., RR-A3299-1, WHAT HAPPENS WHEN JUDGES FOLLOW THE RECOMMENDATIONS OF PRETRIAL DETENTION RISK ASSESSMENT INSTRUMENTS MORE OFTEN? 8 (2024) https://www.rand.org/pubs/research_reports/RRA3299-1.html; Desmarais et al., *supra* note 31; Jodi L. Viljoen et al., *Are risk assessment tools more accurate than unstructured judgments in predicting violent, any, and sexual offending? A meta-analysis of direct comparison studies*, BEHAV. SCI. & LAW (2024), https://onlinelibrary.wiley.com/doi/full/10.1002/bsl.2698; Zhiyuan Lin et al., *The limits of human predictions of recidivism*, 6 SCI. ADV. (2020), https://www.science.org/doi/10.1126/sciadv.aaz0652; R. Karl Hanson & Kelly E. Morton-Bourgon, *The Accuracy of Recidivism Risk Assessments for Sexual Offenders: A Meta-Analysis*, PUBLIC SAFETY AND EMERGENCY PREPAREDNESS CANADA (2007), https://publications.gc.ca/collections/Collection/PS3-1-2007-1E.pdf ; D. A. Andrews et al., *The Recent Past and Near Future of Risk and/or Need Assessment*, 52 CRIME & DELINQ. 7 (2006), https://journals.sagepub.com/doi/10.1177/0011128705281756; D. Mossman, *Assessing Predictions of Violence: Being Accurate about Accuracy*, 62 J. CONSULT. CLIN. PSYCHOL. 783 (1994), https://pubmed.ncbi.nlm.nih.gov/7962882/; Don M. Gottfredson, *Effects of Judges' Sentencing Decisions on Criminal Careers*, National Institute of Justice: Research in Brief (November, 1999),

Modern models that are widely used have been evaluated for differing predictive performance by race or gender, with mixed results.[55] As discussed further below, studies on the predictive performance and disparities of risk assessment tools have significant limitations, and results substantially differ by testing methods, prediction and bias metrics, populations of individuals studied, and translation from quantitative to qualitative findings.

It is possible that use of more advanced machine learning methods could meaningfully improve the performance and bias characteristics of risk assessment tools.[56] This area of research is nascent, however, and advanced models may be more difficult to understand and analyze.

The Risks of Risk Assessment

While there are potential benefits to risk assessment, those benefits may not be fully realized. Use of these tools can also potentially reinforce bias, inequality, and other problems in the criminal justice system. When considering implementation of a risk assessment tool, careful analysis of possible downsides is important.

Accuracy, which is a leading rationale for risk assessment in criminal justice, is also a leading concern. Applying standards widely used in scientific research, validation studies indicate that risk assessment models can better predict criminal justice outcomes than random chance, and possibly in some circumstances better than human judgment alone, but a risk materializing (or not) remains far from certain.[57] As an example, in one state's recent comparative evaluation of several

---

https://www.ojp.gov/pdffiles1/nij/178889.pdf; *see generally* Michael J. White et al, *The Meta-Analysis of Clinical Judgment Project: Fifty-Six Years of Accumulated Research on Clinical Versus Statistical Prediction*, 34 THE COUNSELING PSYCHOL. 341 (2006), https://journals.sagepub.com/doi/10.1177/0011000005285875; William M. Gove et al., *Clinical Versus Mechanical Prediction: A Meta-Analysis*, 12 PSYCHOL. ASSESS. 19 (2000), http://zaldlab.psy.vanderbilt.edu/resources/wmg00pa.pdf; Robyn M. Dawes et al., *Clinical Versus Actuarial Judgment*, 243 SCIENCE 1668 (1989), https://meehl.umn.edu/sites/meehl.umn.edu/files/files/138cstixdawesfaustmeehl.pdf.

[55] See the validation studies for risk assessment tools referenced in the first section of this chapter. There has been limited meta-analysis of relevant research. *See* Sarah L. Desmarais et al., *Predictive Validity of Pretrial Risk Assessments: A Systematic Review of the Literature*, 48 CRIM. JUST. & BEHAV. 398 (2021) (surveying validation studies for pretrial risk assessment tools and noting in supplementary material the demographics in studies); Seena Fazel et al., *The Predictive Performance of Criminal Risk Assessment Tools Used at Sentencing: Systematic Review of Validation Studies*, 81 J. CRIM. JUST., July–Aug. 2022 (surveying validation studies for risk assessment tools used at sentencing and noting population demographics). There has also been limited research directly comparing risk assessment tools using similar populations and methods. *See* JUD. COUNCIL CAL., PRETRIAL RISK ASSESSMENT TOOL VALIDATION (2022), https://www.courts.ca.gov/documents/Pretrial-Pilot-Program-Risk-Assesment-Tool-Validation-2022.pdf (reporting results, including on race and gender disparities, from a coordinated multicounty validation study that examined several risk assessment tools).

[56] *Compare* KLEINBERG ET AL., supra note 33 at 259-260 (finding that gradient-boosted decision trees substantially outperform logistic regression in predicting failure to appear) with Jongbin Jung et al., *Simple Rules for Complex Decisions*, April 2017, https://doi.org/10.48550/arXiv.1702.04690 (finding that simple rules can have equivalent performance to random forest models in predicting failure to appear).

[57] JUD. COUNCIL CAL., PRETRIAL RISK ASSESSMENT TOOL VALIDATION (2022), https://www.courts.ca.gov/documents/Pretrial-Pilot-Program-Risk-Assesment-Tool-Validation-2022.pdf; Sarah L. Desmarais et al., *Performance of Recidivism Risk Assessment Instruments in U.S. Correctional Settings, in* HANDBOOK OF RECIDIVISM RISK/NEEDS ASSESSMENT TOOLS 15 (J.P. Singh, D.G. Kroner, J.S. Wormith, S.L. Desmarais & Z. Hamilton eds., 2018), https://onlinelibrary.wiley.com/doi/abs/10.1002/9781119184256.ch1;

pretrial risk assessment instruments, among individuals with the highest possible risk score, about a sixth to a quarter (depending on the tool) were subsequently arrested for a violent offense.[58] As even this high arrest rate reflects, users of the tool should bear in mind that risk is not inevitability.

Beyond the field of criminal justice, research has demonstrated that even with exceptionally high-quality and long-term data, and even with sophisticated machine learning models, accurately predicting life outcomes such as an eviction or job layoff can be beyond reach.[59] Outcomes in the criminal justice system are similarly products of complex and interrelated individual, environmental, and societal factors.

Perpetuation of bias and inequality is also a significant risk.[60] Studies of risk assessment tools have, in some instances, found notable differences in predictive performance by race and gender.[61] Moreover, a model that has similar predictive performance across groups by some metrics may have substantial differences by others. For example, a risk assessment tool that has comparable precision in predicting recidivism across demographic groups may have very different false positive rates across groups.[62] Researchers have demonstrated that, under realistic assumptions, it can be mathematically impossible to achieve equality across multiple different bias

---

DESMARAIS ET AL (2021) *supra* note 55 at 398; Seena Fazel et al., *The predictive performance of criminal risk assessment tools used at sentencing: Systematic review of validation studies*, 81 J. CRIM. JUST. (2022), https://pmc.ncbi.nlm.nih.gov/articles/PMC9755051/pdf/main.pdf; Anne A. H. de Hond et al., *Interpreting area under the receiver operating characteristic curve*, 4 THE LANCET DIGITAL HEALTH 853 (2022), https://www.thelancet.com/journals/landig/article/PIIS2589-7500(22)00188-1/fulltext; Seth J. Prins & Adam Reich, *Criminogenic risk assessment: A meta-review and critical analysis*, 23 PUNISH. & SOC'Y 578 (2021) https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9385164/; Sonja B. Starr, *Evidence-Based Sentencing and the Scientific Rationalization of Discrimination*, 66 Stan. L. Rev. 803, 842-62 (2014).

[58] JUD. COUNCIL CAL., PRETRIAL RISK ASSESSMENT TOOL VALIDATION (2022), https://www.courts.ca.gov/documents/Pretrial-Pilot-Program-Risk-Assesment-Tool-Validation-2022.pdf.

[59] *E.g.*, Matthew J. Salganik, *Measuring the predictability of life outcomes with a scientific mass collaboration*, 117 PROCS. NAT'L ACAD. SCIS. 8399 (2020), https://www.pnas.org/doi/10.1073/pnas.1915006117.

[60] This discussion focuses on race and gender disparities, which have been a primary focus of relevant research. Other types of disparities may exist and have received limited study, such as on the basis of disability or language. For example, a risk assessment tool that uses employment as a risk factor without considering whether the person receives Social Security Disability Insurance may result in discrimination on the basis of disability. There is also limited research on disparities across subgroups combining race and gender.

[61] *See* JUD. COUNCIL CAL., PRETRIAL RISK ASSESSMENT TOOL VALIDATION (2022), https://www.courts.ca.gov/documents/Pretrial-Pilot-Program-Risk-Assesment-Tool-Validation-2022.pdf; Desmarais et al (2021) *supra* note 55 at 398; Howard Henderson et al., *Determining Racial Equity in Pretrial Risk Assessment*, 86 FED. PROB. 26 (2022), https://www.uscourts.gov/sites/default/files/86_3_5.pdf; Matthew DeMichele et al., *The Public Safety Assessment: A Re-Validation and Assessment of Predictive Utility and Differential Prediction by Race and Gender in Kentucky* (Apr. 25, 2018), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3168452.

[62] *E.g.*, Julia Angwin et al., *Machine Bias*, ProPublica (May 23, 2016), https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing (demonstrating these properties in an evaluation of the COMPAS risk assessment tool).

metrics like these.[63] Navigating the subtle and complex tradeoffs across predictive performance and bias measures is a challenge inherent in risk assessment.[64]

Data is one potential source of bias for risk assessment models. Models are based on data from the criminal justice system, which can encode existing disparities. Consider, as an example, a risk assessment model that uses housing stability as an input and that outputs a score for risk of future arrest. Housing stability may depend on biases in the housing market, and arrests may be affected by biases in policing. The model may incorporate and perpetuate these biases, much like it incorporates other trends in data. Some scholars question the extent to which risk assessment can mitigate disparities in the criminal justice system and society; risk assessment involves prediction based on past events, which may themselves reflect inequality.[65] Some scholars have further suggested that risk assessment may create circular processes, in which individuals or groups are affected by the criminal justice system, then flagged as higher risk, leading to further involvement in the system.[66]

The process of designing, implementing, and validating risk assessment tools may also be affected by disparities. Risk assessment tools may perpetuate inequality when they are developed in a manner that does not incorporate, for example, the perspectives of individuals within affected communities.[67] Tools may also be developed using data that is readily available, rather than alternative types of data that may be more predictive of criminal justice outcomes and less prone to encoding historical biases (e.g., using arrests rather than convictions).[68]

Data quality is another area of potential concern with risk assessment tools. Discrepancies in how data is collected and categorized can result in misleading, incomplete, and inaccurate records that do not accurately reflect factors relevant to risk assessment.[69] Some inputs to risk

---

[63] Jon Kleinberg et al., *Inherent Trade-Offs in the Fair Determination of Risk Scores*, Nov. 2016, https://arxiv.org/abs/1609.05807; Alexandra Chouldechova, *Fair Prediction with Disparate Impact: A Study of Bias in Recidivism Prediction Instruments*, Feb. 2017, at 2, 4, https://arxiv.org/abs/1703.00056; Richard Berk et al., *Fairness in Criminal Justice Risk Assessments: the State of the Art*, 50 Sociol. Methods & Res. 3 (2018), https://journals.sagepub.com/doi/pdf/10.1177/0049124118782533.

[64] *See, e.g.*, Andrew Bell et al., *The Possibility of Fairness: Revisiting the Impossibility Theorem in Practice* (FaccT '23: Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency, 2023), https://dl.acm.org/doi/abs/10.1145/3593013.3594007.

[65] *See, e.g.*, Sandra G. Mayson, *Bias In, Bias Out*, 129 YALE L.J. 2122 (2019), https://www.yalelawjournal.org/article/bias-in-bias-out; Ben Green, *The False Promise of Risk Assessments: Epistemic Reform and the Limits of Fairness*, ACM FAT* (2020), https://dl.acm.org/doi/10.1145/3351095.3372869.

[66] *See, e.g.*, Shawn D. Bushway, *"Nothing Is More Opaque Than Absolute Transparency": The Use of Prior History to Guide Sentencing*, 2.1 HARVARD DATA SCI. REV. (2020), https://hdsr.mitpress.mit.edu/pub/dudgcmk3/release/7.

[67] Ngozi Okidegbe, *The Democratizing Potential of Algorithms?*, 53 CONN. L. REV. 739 (2022), https://scholarship.law.bu.edu/cgi/viewcontent.cgi?article=4138&context=faculty_scholarship.

[68] Jessica M. Eaglin, *Constructing Recidivism Risk*, 67 EMORY L.J. 59, 101-104 (2017), https://scholarlycommons.law.emory.edu/cgi/viewcontent.cgi?article=1046&context=elj.

[69] *See, e.g.*, Sarah Lageson, *Criminally Bad Data: Inaccurate Criminal Records, Data Brokers, and Algorithmic Injustice*, 2023 U. ILL. L. REV. 1771, 1775–81, 1786 (describing sources of criminal record errors and the impact on automated decision-making). For example, inconsistent collection and coding of individuals' race and ethnicity can complicate efforts to understand how justice outcomes for people of color differ from those of white individuals. *See* KELLY ROBERTS FREEMAN, CATHY HU & JESSE JANNETTA, RACIAL EQUITY AND CRIMINAL JUSTICE RISK ASSESSMENT, URBAN INST. 3-4 (2021), https://www.urban.org/sites/default/files/publication/103864/racial-equity-and-criminal-justice-risk-assessment.pdf.

assessment models can be subjective, such as the quality of an individual's relationship with their family, introducing possible inaccuracy and bias.[70] There has been little evaluation about how consistent practitioners and individuals being assessed are in determining these inputs.[71]

Transparency is also a concern. Individuals who are subject to a risk assessment tool (and their representatives) may not know that the tool was used or have sufficient information to understand how it works and how it performs. Affected individuals also may not be aware of the inputs provided to the tool or have an opportunity to correct mistakes. While some models are entirely open, providing free public access to design documentation, implementing materials, and validation studies, others are more restrictive. Commercial licensing requirements, nondisclosure agreements, and trade secret protections, for example, can inhibit evaluation and understanding.[72] The White House Blueprint for an AI Bill of Rights recommends that AI models used in sensitive domains, such as criminal justice, should provide "meaningful access to examine the system."[73] Access to models is essential for enabling independent evaluation to quantify performance and identify issues.

The use of risk assessment models in contexts where they have not been properly validated is another source of concern. A model that is trained on one population, at one time, for one purpose may perform very differently on other populations, at later times, or when used for other purposes.[74] There are significant differences in criminal trends and criminal law across jurisdictions, calling into question how well risk assessment models generalize. Risk assessment models are often used without any recent validation on the local population.[75] It is a best practice

---

[70] *See* Beth Karp, *What Even Is a Criminal Attitude? –And Other Problems with Attitude and Associational Factors in Criminal Risk Assessment*, 75 STAN. L. REV. 1431 (2023), https://review.law.stanford.edu/wp-content/uploads/sites/3/2023/06/Karp-75-Stan.-L.-Rev.-1431.pdf .

[71] Sarah L. Desmarais et al., *Performance of Recidivism Risk Assessment Instruments in U.S. Correctional Settings*, in HANDBOOK OF RECIDIVISM RISK/NEEDS ASSESSMENT TOOLS (J.P. Singh, D.G. Kroner, J.S. Wormith, S.L. Desmarais & Z. Hamilton eds., 2018), https://onlinelibrary.wiley.com/doi/abs/10.1002/9781119184256.ch1.

[72] Hannah Bloch-Wehba, *Access to Algorithms*, 88 FORDHAM L. REV. 1265 (2020), https://ir.lawnet.fordham.edu/cgi/viewcontent.cgi?article=5649&context=flr; Rebecca Wexler, *Life, Liberty, and Trade Secrets: Intellectual Property in the Criminal Justice System*, 70 STAN. L. REV. 1343 (2018), https://review.law.stanford.edu/wp-content/uploads/sites/3/2018/06/70-Stan.-L.-Rev.-1343.pdf ; Cynthia Rudin et al., *The Age of Secrecy and Unfairness in Recidivism Prediction*, 2.1 HARV. DATA SCI. REV. (2020), https://hdsr.mitpress.mit.edu/pub/7z10o269/release/7; Cynthia Rudin et al., *Broader Issues Surrounding Model Transparency in Criminal Justice Risk Scoring*, 2.1 HARV. DATA SCI. REV. (2020), https://hdsr.mitpress.mit.edu/pub/8jy98s9q/release/3; Greg Ridgeway, *Transparency, Statistics, and Justice System Knowledge is Essential for Science of Risk Assessment*, 2.1 HARV. DATA SCI. REV. (2020), https://hdsr.mitpress.mit.edu/pub/vu6rc1yv/release/7.

[73] EXEC. OFF. OF THE PRESIDENT, BLUEPRINT FOR AN AI BILL OF RIGHTS 51 (2022), https://www.whitehouse.gov/ostp/ai-bill-of-rights/ .

[74] ALEXANDRA CHOULDECHOVA & KRISTIAN LUM, SAFETY & JUST. CHALLENGE, THE PRESENT AND FUTURE OF AI IN PRE-TRIAL RISK ASSESSMENT INSTRUMENTS 3, 5 (2020), https://safetyandjusticechallenge.org/resources/the-present-and-future-of-ai-in-pre-trial-risk-assessment-instruments/; Erika Montana et al., *Cohort bias in predictive risk assessments of future criminal justice system involvement*, 120 PROCS. NAT'L ACAD. SCIS. (2023), https://www.pnas.org/doi/10.1073/pnas.2301990120.

[75] *50-State Report on Public Safety Part 2, Strategy 2, Action Item 2*, THE COUNCIL OF STATE GOVERNMENTS, https://50statespublicsafety.us/part-2/strategy-2/action-item-2/.

to build a predictive model on the most representative data available, or failing that, to evaluate a model on representative data to determine if it is suitable for use.[76]

There is limited research on the effects of risk assessment tools, both on the criminal justice systems and officials who use them and on the communities where they are implemented.[77] More research is needed to understand how practitioners integrate risk assessment information into decision-making, including to examine the risks of automation and confirmation biases, why officials sometimes choose to override recommendations based on risk assessment, and how risk assessment affects the accuracy and disparities of decisions.[78] Similarly, more research is needed on how risk assessment can impact communities. Studies of this type typically lack adequate controls to disentangle implementation of risk assessment from other trends, and the time horizon for studies may be too short to capture benefits and downsides.

Converting model output into categories (e.g., "high risk"), rather than presenting probabilities and confidence intervals, may omit important information for decision-makers.[79] An individual may be just over or under the threshold for a risk category, for example. While risk categories may have value in helping to explain results and counter precision bias, omitting more detailed information can deprive decision-makers of essential context.

A final challenge is that evaluating risk is fundamentally different from determining the appropriate treatment for an individual from a range of possible treatments.[80] Focusing on prediction of negative outcomes such as recidivism or failure to appear could divert resources from, and limit consideration of, interventions that may reduce the risk of those negative outcomes

---

[76] For example, the Minnesota Department of Corrections compared a nationally available tool to a state-specific one and concluded that "there is a home-field advantage to risk assessment," because the Minnesota tool outperformed an off-the-shelf tool. Grant Duwe, Mn. Dep't of Corr., Evaluating Bias, Shrinkage and the Home-Field Advantage: Results from a Revalidation of the MnSTARR 2.0, at 29–31 (2021).

[77] Jodi L. Viljoen et al., *Impact of Risk Assessment Instruments on Rates of Pretrial Detention, Postconviction Placements, and Release: A Systematic Review and Meta-Analysis*, 43 L. HUM. BEHAV. 397 (2019), https://psycnet.apa.org/fulltext/2019-46921-001.pdf; Megan Stevenson, *Assessing Risk Assessment in Action*, 103 MINN. L. REV. 303 (2018), https://scholarship.law.umn.edu/cgi/viewcontent.cgi?article=1057&context=mlr . Shamena Anwar et al., RAND, RR-A3299-1, WHAT HAPPENS WHEN JUDGES FOLLOW THE RECOMMENDATIONS OF PRETRIAL DETENTION RISK ASSESSMENT INSTRUMENTS MORE OFTEN? 8 (2024), https://www.rand.org/pubs/research_reports/RRA3299-1.html ; Matthew DeMichele et al., *What Do Criminal Justice Professionals Think About Risk Assessment at Pretrial?*, 83 FED. PROB. 32 (2019). https://www.uscourts.gov/sites/default/files/83_1_5_0.pdf; Sarah Riley, *Overriding (In)justice: Pretrial Risk Assessment Administration on the Frontlines*, FaccT '24: Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency, 2024), https://facctconference.org/static/papers24/facct24-35.pdf; Brandon Garrett & John Monahan, *Judging Risk*, 108 CAL. L. REV. 439 (2020). https://www.californialawreview.org/print/judging-risk ; Megan T. Stevenson & Jennifer L. Doleac, *Algorithmic Risk Assessment in the Hands of Humans*, AM. ECON. J.: ECON. POL'Y. (Nov. 2024), https://www.aeaweb.org/articles?id=10.1257/pol.20220620.

[79] Melissa Hamilton, *Risk Assessment Tools in the Criminal Legal System – Theory and Practice: A Resource Guide*, NAT'L ASS'N OF CRIM. DEF. LAWS. 44-48 (Nov. 2020), https://www.nacdl.org/getattachment/a92d7c30-32d4-4b49-9c57-6c14ed0b9894/riskassessmentreportnovember182020.pdf ; Starr (2014), *supra* note 57.

[80] Chelsea Barabas et al., *Interventions over Predictions: Reframing the Ethical Debate for Actuarial Risk Assessment* (Conference on Fairness, Accountability, and Transparency, Proceedings of Machine Learning Research, 2018), https://proceedings.mlr.press/v81/barabas18a/barabas18a.pdf; Erin Collins, *Punishing Risk*, 107 GEO. L.J. 57 (2018), https://www.law.georgetown.edu/georgetown-law-journal/wp-content/uploads/sites/26/2018/12/Punishing-Risk-2.pdf; Starr (2014), *supra* note 57.

and otherwise improve individual rehabilitation and public safety. For example, emphasizing criminal justice risks may deemphasize interventions outside the criminal justice system, such as substance abuse treatment or job training, that could have positive effects.

Recommendations

a. Risk Assessment Tool Design, Implementation, and Validation

Before moving forward with a risk assessment tool, it is important to document objectives, performance and bias criteria, expected benefits and risks, and alternatives to using the tool. Risk assessment tools should only be used if the expected benefits clearly outweigh the risks, substantiated by adequate evidence. In some circumstances, it may be preferable to use trained professionals and structured judgments rather than an actuarial risk assessment model.

Throughout the process of deciding whether and how to use risk assessment, engagement with affected communities and stakeholders is essential. Their perspectives are vital for setting goals, identifying possible issues, and integrating mitigation strategies. Because risk assessment tools are built and validated on data, a critical early step is careful evaluation of the quality of available data, identifying gaps, inconsistencies, or biases that could affect risk assessment. Agencies should consider setting policies and procedures in place to collect, label, and store data efficiently and accurately, ensuring the integrity, completeness, and provenance of the data. Transparency about available data is a valuable part of maximizing the quality of the data used, as stakeholders may be well-positioned to identify opportunities and shortcomings with that data. If a tool involves risk factors within an individual's control, transparency can also enable affected individuals to better understand how to reduce their risk score.

If data under consideration for a risk assessment tool includes information covered by antidiscrimination law (e.g., race, gender, disability, or age), or foreseeable proxies for those protected characteristics (e.g., ZIP code as a proxy for race), designers and implementers should be especially cautious and ensure legal compliance. They should also bear in mind that large criminal justice databases are almost certain to contain errors. Criminal justice data may also reflect inconsistencies in recordkeeping and enforcement. Developers of risk assessment tools should account for these issues, trying to quantify them and build tools that are robust against errors and inconsistencies. At the same time, if there are types of data available that can serve a similar purpose in risk assessment, they should carefully consider tradeoffs. Relying on charges filed or convictions rather than arrests may reduce the possibility of bias, for example, by accounting for disparities in policing and involving additional components of the criminal justice system.

Developers of risk assessment tools should also be mindful of the value of a fit between the data they are utilizing and the jurisdiction in which the tool will be used. The data used for the output of the risk assessment tool should also closely resemble the decision that the tool will support. For example, a tool for predicting misconduct in custody would ideally by built with data about misconduct in custody, rather than data about recidivism after release. Also, where there is sufficient data in a jurisdiction, it is preferable to build a model specific to that jurisdiction rather than rely on a model built with data that may not be representative. And if a risk assessment tool

uses data from multiple jurisdictions, the tool should account for possible differences in data definitions across jurisdictions.

Risk assessment tools should also be evaluated in a real-world context before deployment. First, this evaluation should include performance in predicting real-world outcomes, on the specific population whose risk is being assessed, as well as measurement of predictive performance across demographic groups, to identify possible biases. There should also be a baseline included in the evaluation, such as predictions made by a human or a previous model, to comparatively assess a tool's performance. Second, deployment of new risk assessment tools should be phased, facilitating observation and mitigation of risks. This approach also provides data for comparatively assessing the community impacts of using risk assessment tool. Third, the design, implementation, and validation of a risk assessment tool should be reviewed by an evaluator who is independent of the tool's developer. Finally, if significant risks are identified before deployment, mitigations for those risks should be designed, implemented, and validated before the deployment continues.

b. Continuous Monitoring and Human Oversight

Continuous monitoring of risk assessment tools is essential for ensuring they continue to meet objectives and rapidly identifying emerging issues. These tools should be continuously monitored to identify possible changes in predictive performance, biases, or data. This monitoring should be automated, to the extent possible, to allow quick responses to changes. Tools should also be periodically revalidated, to ensure that trends in criminal justice and changes in laws, policies, and practices have not undermined performance. Similarly, tools should be retrained or updated on a periodic basis, to account for changing context. This process should consider prior feedback and outcomes, which may suggest ways of improving the tool's design.

Risk assessment tools should be paired with policies and procedures, and ideally automated monitoring, to ensure the accuracy of input data. They should not displace human decision-making in the criminal justice system. They should be accompanied by guidance and training for decision-makers about the tool and its limitations.

Individuals affected by a risk assessment tool should receive notice and explanation and should have an opportunity to respond. When a decision-maker in the criminal justice system uses a risk assessment tool, to the extent possible, the individual being assessed (and their representatives) should receive notice of the tool and explanation of how it is being used. Where feasible, an individual should be able to seek correction of errors in any input data or submit additional relevant context.

c. Training and Education

Individuals who use risk assessments in criminal justice should be trained on how to interpret a tool's scores, as well as the limits of prediction. With respect to limits, training curricula should pay special attention to the potential for bias and creating or perpetuating disparities.

Users of a risk assessment tool, including judges, prosecutors, defense attorneys, pretrial or probation officers, correctional staff, community supervision staff, and others, should be

informed about how the tool works and its limitations. Training programs and other educational offerings should evolve with changes in research and the tool.

d. Policies

Agencies should issue guidance to risk assessment users on appropriate and prohibited uses of the risk assessment tool. Agencies should also issue guidance to risk assessment users on when decisions from the risk assessment tool may require additional human oversight and intervention—for example, when there is high uncertainty in a risk assessment tool's predictions, or when a decision has particularly high impact.

Agencies should provide public notice and documentation on the use of the system, as well as information on its design (including input data) and the testing conducted to mitigate risks of the system.

Agencies should ensure that there is meaningful access to the risk assessment tool, such that researchers and stakeholders can conduct their own evaluations.

Agencies should only use risk assessment tools where the prediction model can be made transparent to the public. Agencies should not use risk assessment tools where this core functionality is obscured by trade secret protections or other barriers.

Agencies should not use tools with prediction thresholds that cannot be evaluated and changed by agency policymakers.

Agencies should set a model's prediction thresholds according to a methodology for achieving specific objectives, and agencies should provide decision-makers with individualized context about a risk assessment beyond a categorization.

e. Research

Future research should examine risk assessment tools to better understand predictive performance, biases, how they compare to alternatives, and how they affect decision-makers and communities.

VI. Conclusion & Best Practices

The preceding chapters offer insight into uses of AI within the criminal justice system. These uses potentially offer important benefits to the institutions and individuals in the criminal justice system—including judges, pretrial and probation officers, prosecutors, law enforcement officers, forensic professionals, criminal defendants, and defense attorneys—and to the broader public. The preceding chapters also discuss significant risks as well as important policy and technological considerations that users should understand and mitigate. Each chapter underscores that responsible use of AI is critical, especially in the criminal justice context, where the deployment of this technology is quickly evolving and where public safety and individual rights are on the line. Each chapter also recommends practices to mitigate risks and preserve privacy, civil liberties, and civil rights when using AI in the criminal justice context.

Central to all these recommendations is AI governance. As a general matter, governance "provides a framework for decision-making by establishing standards and procedures and clarifying roles and responsibilities."[1] As applied to AI, governance includes the broad range of decisions surrounding the technology's development, use, and safeguards. Further, in the context of the criminal justice system, AI governance must be adaptable to account for change, but it must also be grounded in enduring values. Indeed, AI governance in this space must account for civil rights and civil liberties just as much as technical considerations such as data quality and data security. Governance is central to the Department of Justice's responsible use of AI, and this report emphasizes the importance of AI governance at the diverse law enforcement agencies and other criminal justice institutions using AI or contemplating its use.

AI governance will take different shapes, depending on the size, scope, mission, and resources of each actor or institution operating within the criminal justice system. As with other forms of governance, smaller or resource-constrained criminal justice actors—including those at the state, local, and municipal level—may not be in a position to fully implement each recommendation exactly as written in this report. However, a key takeaway is that AI presents a new set of considerations that actors and institutions in the criminal justice system should take into account in adapting their overall governance structures.

The recommendations included below, and detailed in each of the preceding chapters, draw from the work of a community of scholars across the United States and abroad, as well as the work of our colleagues elsewhere in the Federal Government on publications such as:

- The White House Office of Science and Technology Policy's "The Blueprint for an AI Bill of Rights: Making Automated Systems Work for the American People"[2];

[1] Governance, Digital.gov, https://digital.gov/topics/governance/.
[2] EXEC. OFF. OF THE PRESIDENT, BLUEPRINT FOR AN AI BILL OF RIGHTS 5 (Oct. 2022), https://www.whitehouse.gov/wp-content/uploads/2022/10/Blueprint-for-an-AI-Bill-of-Rights.pdf.

- NIST's "Artificial Intelligence Risk Management Framework,"[3] and Special Publication 1270, "Towards a Standard for Identifying and Managing Bias in Artificial Intelligence"[4];
- OMB's Memorandum M-24-10, "Advancing Governance, Innovation, and Risk Management for Agency Use of Artificial Intelligence,"[5] and Memorandum M-24-18, "Advancing the Responsible Acquisition of Artificial Intelligence in Government"[6];
- "National Security Memorandum on Advancing the United States' Leadership in Artificial Intelligence; Harnessing Artificial Intelligence to Fulfill National Security Objectives; and Fostering the Safety, Security, and Trustworthiness of Artificial Intelligence"[7] and the associated "Framework to Advance AI Governance and Risk Management in National Security."[8]

The recommendations in this chapter are also goals that the Department of Justice is working toward for its own uses of AI, as described in the Department's "Compliance Plan for OMB Memorandum M-24-10."[9]

Foundations for AI Governance

To establish a durable and comprehensive AI governance program, criminal justice agencies should identify the problem to solve and the reasons why the use of AI is preferable to alternatives; establish clear organizational and reporting structures to provide oversight, monitoring, and evaluation; hire, train, and retain a workforce with adequate resources to devise, enact, and enforce policies; build, operate, and routinely monitor the AI systems; regularly evaluate the performance of systems for accuracy and any unintended biases or disparities; and mitigate their associated risks.

---

[3] NAT'L INST. STANDARDS & TECH., ARTIFICIAL INTELLIGENCE RISK MANAGEMENT FRAMEWORK: GENERATIVE ARTIFICIAL INTELLIGENCE PROFILE (July 2024), available at https://doi.org/10.6028/NIST.AI.600-1.

[4] Reva Schwartz et al., NAT. INST. STANDARDS & TECH. SPECIAL PUBL'N 1270, TOWARDS A STANDARD FOR IDENTIFYING AND MANAGING BIAS IN ARTIFICIAL INTELLIGENCE (2022), https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.1270.pdf.

[5] MEMORANDUM FROM SHALANDA D. YOUNG, DIR., OFF. MGMT. & BUDGET, EXEC. OFF. OF THE PRESIDENT, TO HEADS OF EXEC. DEP'TS & AGENCIES, (Mar. 28, 2024), available at https://www.whitehouse.gov/wp-content/uploads/2024/03/M-24-10-Advancing-Governance-Innovation-and-Risk-Management-for-Agency-Use-of-Artificial-Intelligence.pdf.

[6] MEMORANDUM FROM SHALANDA D. YOUNG, DIR., OFF. MGMT. & BUDGET, EXEC. OFF. OF THE PRESIDENT, TO HEADS OF EXEC. DEP'TS & AGENCIES, (Sep. 24, 2024), available at https://www.whitehouse.gov/wp-content/uploads/2024/10/M-24-18-AI-Acquisition-Memorandum.pdf.

[7] MEMORANDUM ON ADVANCING THE UNITED STATES' LEADERSHIP IN ARTIFICIAL INTELLIGENCE; HARNESSING ARTIFICIAL INTELLIGENCE TO FULFILL NATIONAL SECURITY OBJECTIVES; AND FOSTERING THE SAFETY, SECURITY, AND TRUSTWORTHINESS OF ARTIFICIAL INTELLIGENCE (Oct. 24, 2024), available at https://www.govinfo.gov/app/details/DCPD-202400945.

[8] FRAMEWORK TO ADVANCE AI GOVERNANCE AND RISK MANAGEMENT IN NATIONAL SECURITY (OCT. 24, 2024), available at https://ai.gov/wp-content/uploads/2024/10/NSM-Framework-to-Advance-AI-Governance-and-Risk-Management-in-National-Security.pdf.

[9] U.S. DEP'T OF JUST., COMPLIANCE PLAN FOR OMB MEMORANDUM M-24-10 (Oct. 2024), available at https://www.justice.gov/media/1373026/dl.

a. Conduct a Needs and Alternatives Assessment

Criminal justice agencies should know what problem or function they are trying to solve or address and how both AI and non-AI alternatives might present a solution.[10] For example, judges, pretrial and probation officers, prosecutors, law enforcements officers, and defense attorneys may be seeking to more accurately calculate the likelihood that an individual will have a particular outcome in the criminal justice system. That calculation can in turn help judges make transparent, non-discriminatory, and equitable decisions. Similarly, agencies seek to drive law enforcement efficiencies and best direct limited resources. Agencies could consider risk assessment tools as compared to subjective assessments, or predictive policing tools as compared to traditional resource allocation measures.

b. Establish a Clear Organizational Structure

Criminal justice agencies, like other institutions working within the criminal justice system, vary in size, scope, resources, and mission. Regardless of institutional size, however, a tiered human-centered structure that assigns to individuals distinct operational, decision-making, and oversight responsibilities commensurate with their function, training, and experience will be key to maintaining an effective AI governance program. For example, at the Department of Justice, the Attorney General has designated a Chief AI Officer (CAIO) and an Emerging Technology Board (ETB), as required in OMB M-24-10. The CAIO has primary responsibility for coordination of the Department's use of AI, efforts to promote AI innovation, and management of risks from the use of AI. The CAIO coordinates with components across DOJ and reports to the Deputy Attorney General. The ETB serves as the Department's AI Governance Board.

c. Catalog Uses of AI

Agencies should catalog existing and new uses of AI and facilitate organization-wide visibility into these uses. Awareness of the AI uses occurring within the agency is the foundation for other governance steps. As feasible, agencies should also aim for greater transparency by making their use cases public. For example, the Department of Justice's AI governance process begins with identifying existing, new, and planned uses of AI across the Department. Components will report AI use cases, per the Advancing American AI Act and OMB M-24-10, and the Department will also review procurement, privacy governance, and IT governance records to ensure comprehensiveness.

d. Hire, Train, and Retain AI Workforce

Agencies interested in using AI should hire, train, and retain a workforce that understands the technical aspects, operational considerations, and tradeoffs associated with the use of AI (and remains knowledgeable as AI rapidly evolves). An adequately trained workforce should be able to understand, among other things, the source and limitations of the data on which a system is trained, the potential sources of any discrimination and bias, how these sources can affect outcomes in the

[10] For institutions already using AI, this might take the form of an inventory that addresses the functionality of the tool (what does it do?), alternative tools (are there other options to achieve that result?), and a justification for using the tool that accounts for risks and mitigation strategies (why was the tool chosen among the existing alternatives?).

criminal justice system, the steps necessary to mitigate bias and discrimination, and means of implementing effective monitoring and auditing of AI systems.

Pre-Deployment Measures

Prior to deploying AI systems, law enforcement agencies and other agencies involved in criminal justice should also undertake the following measures.

e. Implement Policies and Procedures

Use of AI in the criminal justice system must be governed by robust human-centered policies and procedures that describe the permitted and prohibited uses of the AI system, the data used to train the system, the metrics and processes used to validate or evaluate the accuracy of the output, the frequency of monitoring, and the risks of using a particular AI system. The use of AI in the criminal justice system must also comply with the Constitution; statutes; rules of evidence, discovery, and procedure; discovery obligations; and ethical commitments. These policies and procedures should be developed by agencies in collaboration with relevant stakeholders and should anticipate and incorporate civil rights and civil liberties concerns. In addition, these policies should be regularly re-evaluated to ensure accuracy and consistency with current technology, best practices, and legal requirements.

Keeping "Humans in the Loop:" Individuals and institutions operating in the criminal justice system that use AI systems must ensure that human judgment drives the system's design, implementation, and use. Users should be mindful that AI cannot displace human decision-making (and, indeed, doing so would be counter to established laws, norms, and principles designating different roles to, for instance, judges, officers, and prosecutors), and AI system outputs should be reviewed and verified by a human being. Courts, agencies, and decision-makers should create decision-making processes and guidelines for the use of AI to ensure a human reviews any outputs for accuracy. For high-impact decisions (such as cause for arrest or length of a sentence), the output of an AI system should not be the sole basis for a decision. For example, Department of Justice decisions regarding investigative steps, detention, prosecution, and evaluation or analysis of forensic evidence may take into account AI outputs to assist in collection and analysis of information, but trained professionals make the decisions.

Community Outreach and Public Notice: Consistent with applicable law and government guidance, and to the extent feasible, criminal justice entities should actively engage the public and relevant stakeholders regarding the intended use of potentially rights-or-safety-impacting AI technologies.[11] Disclosures should not include sensitive operational and law enforcement information but should generally include a plain language description of the system; the data used to train the system (e.g., driver license photos in the case of a facial recognition AI system); who

[11] OMB Memorandum M-24-10 offers useful guidance to agencies assessing whether AI may be rights-impacting or safety-impacting. In particular, M-24-10 defines these terms (note that rights-impacting AI includes AI that, among other factors, may have a significant effect on civil rights, civil liberties, or privacy) and specifies AI uses presumed to be rights- or safety-impacting. An agency's engagement with the community and stakeholders may identify additional, potentially community-specific factors to consider in assessing the possible impact of an AI technology.

will have access to the system; the circumstances for the system's deployment; and governance mechanisms in place, such as whether the system will have an opt-out mechanism once deployed. Disclosure and engagement regarding the general purpose of an AI system might be feasible and useful even if disclosing details of capabilities and implementation is infeasible. Engagement can help build trust and ensure public support for adopted technologies.

Defining and Categorizing Risk: Criminal justice entities using AI should define and measure AI risks to rights and safety; categorize AI uses based on those risks; adopt commensurately greater safeguards for uses with rights- and safety-impacting risks; and determine their risk tolerance as an organization. The potential impact of AI uses varies dramatically, and the rigor of governance processes to manage risk may vary accordingly.

Regular Updates: Policies and procedures will inevitably require updates and should contemplate the circumstances that may warrant revisions. Agencies may consider, for instance, reviewing their AI governance policies on a yearly basis to account for technological and legal developments in this area. Agencies should consult with their legal counsel as needed to make appropriate revisions to reflect intervening legal decisions or changes in the applicable rules or statutes in the civil or criminal context.

f. Identify and Analyze Training Data Quality and Sources

Prior to deploying any AI system, individuals and institutions operating in the criminal justice system should closely track and identify the quality and sources of the data on which the AI system was trained and assess the relevance of the data the system was trained on to their particular use and jurisdiction. This is a crucial consideration for all AI systems but may be particularly relevant to institutions intending to contract AI systems designed, administered, or maintained by third-party vendors who may use proprietary training data sets or may otherwise be unwilling to disclose such information to their clients. To that end, criminal justice actors should obtain legal advice and involve procurement professionals to understand and negotiate contracts that allow those actors to examine the training data and address other important civil rights, civil liberties, or privacy concerns.

g. Test Systems Under Deployment Conditions

Criminal justice entities should conduct rigorous pre-deployment testing of any AI or automated systems. To the extent possible, such testing should be performed by independent third parties following domain-specific best practices and under conditions that mirror actual end-to-end deployment conditions. This testing must account for and seek to reflect the complex operational and societal contexts in which these systems are used. Testing a system in isolation may not accurately or fully reflect the system's impact. Agencies should also use pilots and limited releases in advance of full deployment to identify and mitigate against potential risks. The performance of the AI system should be evaluated against the performance of the status quo—such as technological tools or human processes that AI will be replacing or supplementing—and alternatives, and performance factors should include testing for bias, discrimination, and disparities. To instill public confidence, agencies may also consider facilitating external testing as feasible.

h. Evaluate Risks and Implement Risk Mitigation Strategies

After testing, criminal justice agencies should identify rights- or safety-impacting use cases and possible mitigation measures. Mitigation measures should include reconsidering use of rights- or safety-impacting use cases for which risks cannot be effectively mitigated. The decision not to proceed with a use case must be a feasible option.

i. Create Technology-Specific Policies

Some AI use cases, such as facial recognition, create unique or significant considerations that may justify dedicated policies. Agencies should consider whether specific policies are appropriate for particular use cases.

Post-Deployment Measures

After deploying AI systems, criminal justice agencies should also undertake the following measures.

j. Monitor the AI System

Following deployment, entities should conduct regular audits and monitoring of their uses of AI. Entities should adopt a mechanism for objective, independent auditing of models and source code to proactively address concerns about model accuracy, reliability, outdated training or test data, and potential for bias and discrimination.

k. Evaluate New Uses

Agencies should also consider whether and how their use of the AI system might have evolved over time, or how the context in which the system was deployed may have changed. If agencies identify new uses or new data sources that were not considered or evaluated pre-deployment, or if the context changes substantially, those agencies should implement pre-deployment strategies to the new uses and new data and evaluate whether those new uses and new data present new risks to the rights or safety of individuals, whether the new use is appropriate in light of those risks, and any mitigation strategies to be executed.

l. Continue Engaging with the Public

Post-deployment, agencies should continue to actively consult the public, and in particular, communities they judge will be most likely to be affected by the implementation of AI in the criminal justice system. Community engagement should be proactive, and feedback should be solicited on a regular, ongoing basis. Entities should establish channels for affected community members to seek recourse and alternatives to AI systems where practicable.

While this report considers many critical uses of AI in the criminal justice system as specified by EO 14110, existing and possible uses extend well beyond these cases. Use cases associated with future and emerging AI technologies—such as generative AI—might also have a transformative impact on the criminal justice system. Establishing robust AI governance programs

at criminal justice agencies can position those agencies to address risks of these use cases and better capture their promise.

Legal Disclaimer

This report is intended to support the development of policies and practices that protect civil rights, civil liberties, privacy, and equity, and that promote democratic values in the building, deployment, and governance of automated systems. It reflects and describes academic studies, policy proposals, and advocacy positions that might not be adopted or endorsed by the U.S. government.

This report is non-binding and does not constitute U.S. government policy. It does not supersede, modify, or direct an interpretation of any existing statute, regulation, policy, or international instrument. It does not constitute binding guidance for the public or federal agencies and therefore does not require compliance with the principles described herein. It also is not determinative of what the U.S. government's position will be in any international negotiation. The Department's inclusion of recommendations does not indicate a determination that adoption of those recommendations would necessarily satisfy requirements set forth in existing statutes, regulations, policies, or international instruments, or the requirements of the federal agencies that enforce them. Recommendations are not intended to, and do not, prohibit or limit any lawful activity of a government agency, including law enforcement, national security, or intelligence activities. The appropriate application of the recommendations set forth in this report depends significantly on the context in which automated systems are being utilized. In some circumstances, application of these recommendations in whole or in part may not be appropriate given the intended use of automated systems to achieve government agency missions. Future sector-specific guidance will likely be necessary and important for guiding the use of automated systems in certain settings.

This report recognizes that law enforcement activities require a balancing of equities, for example, between the protection of sensitive law enforcement information and public access to information; as such, public access to information may not be appropriate, or may need to be adjusted to protect sources, methods, and other law enforcement equities. In all circumstances, federal departments and agencies remain subject to judicial and congressional oversight, as well as oversight by private and public civil liberties and privacy-focused organizations, as well as existing laws, policies, and safeguards that govern automated systems.

This report is not intended to, and does not, create any legal right, benefit, or defense, substantive or procedural, enforceable at law or in equity by any party against the United States, its departments, agencies, or entities, its officers, employees, or agents, or any other person, nor does it constitute a waiver of sovereign immunity.

NLC Councils, Groups & Committees Membership Applications

Applications are now open for National League of Cities (NLC) Federal Advocacy Committees and Member Councils. Join a Federal Advocacy Committee to bring your voice to the national conversation or join a Member Council and connect with municipalities like yours. Click here to apply or for additional information. Applicants must be from an NLC member city, town or village – deadline December 2. For questions or to learn more about joining NLC, please e-mail membership@nlc.org.

NLC Federal Advocacy Committees include:

- Community & Economic Development
- Energy, Environment & Natural Resources
- Finance, Administration & Intergovernmental Relations
- Human Development
- Information Technology & Communications
- Public Safety & Crime Prevention
- Transportation & Infrastructure Services

https://www.nlc.org/advocacy/federal-advocacy-committees/

NLC Member Councils include:

- First Tier Suburbs Council
- Large Cities Council
- Military Communities Council
- Small Cities Council
- University Communities Council
- Council on Youth, Education, and Families
- Race, Equity and Leadership (REAL) Council

PSCP Sessions of Interest

Wednesday, Nov 19, 2025

NLCU: AI for City Leaders | Turning Strategy into Action

08:00 AM EST - 08:50 AM EST Salt Palace Convention Center

Ready to harness the power of artificial intelligence to advance your city's goals? Join Bloomberg Philanthropies' What Works Cities for an interactive, hands-on workshop designed specifically for government leaders. This session demystifies AI and shows how it can be a powerful tool to accelerate strategic priorities—from improving service delivery to enhancing community engagement. Participants will explore the essential data foundations for successful AI adoption, including boosting data quality, setting clear standards, and building trust through community involvement. Walk away with practical insights and a roadmap to begin your AI journey with confidence and clarity.

Thursday, November 20, 2025

Drones in Action: Practical Innovation for Cities & Towns

10:00 AM EST - 10:45 AM EST Salt Palace Convention Center

Cities across the country are exploring new tools to strengthen public safety while protecting the rights and trust of their residents. In this fireside chat, join Flock Safety and Los Angeles Councilmember John Lee to discuss how cities are adopting technologies such as license plate recognition, gunshot detection, drones, and live video. The conversation will highlight how these tools are helping communities prevent crime and solve cases more quickly, as well as the policies, transparency practices, and community engagement efforts needed to ensure technology serves residents responsibly. City leaders will walk away with insights on balancing innovation with accountability to create safer, more connected, and more trusted communities.

FirstNet in Action: Enhancing Communication for Safer Communities

11:00 AM EST - 11:45 AM EST Salt Palace Convention Center

Created by Congress after 9/11, FirstNet is transforming how public safety agencies communicate—especially during large-scale events and emergency responses. This session explores how FirstNet's mission-critical broadband solutions integrate with traditional land mobile radio (LMR) systems to deliver greater reach, redundancy, and functionality. Municipal and public safety leaders will share how FirstNet enables unified, on- and off-scene coordination, improves interoperability, and unlocks advanced capabilities—all while driving cost efficiencies. Discover

how cities can strengthen their emergency response infrastructure and equip first responders with the tools they need to protect communities more effectively.

Unlocking City Potential with Geospatial AI

11:00 AM EST - 11:45 AM EST Salt Palace Convention Center

Artificial intelligence (AI) is transforming how cities address their challenges, from day-to-day workflows to their most pressing issues. When paired with the GIS technology that you already use, geospatial AI becomes an intelligent partner — streamlining workflows, revealing hidden patterns and anticipating emerging needs. This session will highlight how city leaders can utilize geospatial AI to manage traffic flow, inform land use decisions, monitor environmental changes and enhance resilience. By automating routine tasks and improving insight across planning, housing, infrastructure and public safety, geospatial AI equips local leaders to act with greater confidence, precision and vision. Join us to see how this next-generation approach is already shaping stronger, more sustainable communities.

Practical Solutions for Communities Using AI

04:00 PM EST - 05:30 PM EST Salt Palace Convention Center

Artificial intelligence (AI) isn't just for tech giants and big cities anymore. As AI tools become more accessible and user-friendly, small and mid-size communities are stepping into the future—leveraging AI to streamline operations, enhance public services, and better meet the needs of their residents. Join this dynamic panel of local government leaders and AI experts as they share real-world stories of innovation from local government. You'll gain practical insights into how AI is being implemented on the ground, with a focus on equity, strategy, and lessons learned along the way. Stick around for our interactive "Office Hours" segment, where you'll have the chance to ask questions, explore use cases, and workshop your own AI challenges with the panel and fellow attendees.

Public Health & Safety Strategies to Prevent Violence

04:00 PM EST - 05:30 PM EST Salt Palace Convention Center

Community violence is not inevitable — it's preventable. Across the country, local leaders are embracing innovative approaches to build safer communities. This workshop will explore a range of effective strategies for reducing and preventing community violence. Participants will examine violence prevention and intervention efforts, including community-based violence interrupter programs. The session will also highlight trauma-informed policing and alternative crisis response models that build trust. Learn the importance of cross-sector partnerships with schools, healthcare providers and grassroots organizations to create comprehensive and sustainable violence prevention strategies.

Friday, November 21, 2025

Designing for a Cooler Future: How Cities Are Tackling Extreme Heat

10:30 AM EST - 12:00 PM EST Salt Palace Convention Center

Extreme heat is becoming one of the most urgent climate threats facing cities—and local leaders are stepping up with innovative, cost-effective solutions to protect their communities. This workshop dives into how municipalities are using nature-based strategies and reimagining built infrastructure to reduce heat risks and improve public health. Hear directly from city leaders about the tools they're developing, the challenges they're overcoming, and how they're turning heat resilience plans into real-world action.

Trust + Safety: Local Government Strategies on Federal Immigration Policy

02:15 PM EST - 03:45 PM EST Salt Palace Convention Center

As immigration policy continues to evolve, local governments are on the front lines—working to balance public safety, legal responsibility and community trust. This timely workshop is designed for city leaders navigating the complex intersection of federal immigration enforcement and local governance. Join legal experts and municipal officials as they share real-world strategies on messaging and coalition building to ensure public health, safety and economic resilience while responding to increased federal immigration enforcement efforts. Learn how local governments are navigating the intersection of federal immigration policy and public health and safety.

Saturday, Nov 22, 2025

Building a Better System to Respond to Homelessness

10:45 AM EST - 12:15 PM EST Salt Palace Convention Center

771,480 people without housing. One national crisis. Countless local solutions. As unsheltered homelessness rises across the country, communities are being called to act with urgency and innovation. This workshop offers a deep dive into how local governments are reimagining their homelessness response systems — from prevention to permanent housing. Participants will explore how cities are preventing homelessness through eviction diversion and reentry planning; partnering with nonprofits and other organizations to align resources and strategies; designing housing continuums that meet the evolving needs of residents; and integrating services, crisis response and long-term solutions. Through real world examples and practical insights, this session will highlight how communities are building comprehensive systems that address both the root causes and immediate impacts of homelessness.

Annual Business Meeting

02:30 PM EST - 04:30 PM EST Salt Palace Convention Center

Schedule At-A-Glance

All events are listed in Mountain Standard Time (MST). Dates and times are subject to change.

TUESDAY, NOVEMBER 18, 2025

Registration & Ask NLC Open (3 p.m. – 5 p.m.)

WEDNESDAY, NOVEMBER 19, 2025

Registration & Ask NLC Open (7 a.m. – 6:30 p.m.)
NLC University Sessions (8 a.m. – 5 p.m.)
Host City Mobile Workshops (9 a.m. – 12 p.m.)
First-Time Attendee Lunch (12 p.m. – 1 p.m.)
Host City Mobile Workshops (1 p.m. – 4 p.m.)
Board Finance, Legislative, and Membership Committee Meetings (1:15 p.m. – 2:30 p.m.)
Advisory Council Meeting (1:30 p.m. – 3 p.m.)
Solution Sessions (2 p.m. – 4:45 p.m.)
Federal Advocacy Committee Meetings (3 p.m. – 5 p.m.)
*Expo Hall Welcome Reception (5 p.m. – 7 p.m.)
State League Directors and Staff Networking Reception (5:00 p.m. – 7:00 p.m.)
Local Indigenous Leaders (LIL) Membership Meeting (5:15 p.m. – 6:15 p.m.)
Hispanic Elected Local Officials (HELO) Membership Meeting (5:15 p.m. – 6:15 p.m.)
Board of Directors Dinner (Invite only) (6 p.m. – 9 p.m.)

THURSDAY, NOVEMBER 20, 2025

Registration & Ask NLC Open (7 a.m. – 5:30 p.m.)
NLC Board of Directors Meeting (8 a.m. – 10 a.m.)
Solution Sessions (9 a.m. – 11:45 a.m.)
Host City Mobile Workshops (9 a.m. – 12 p.m.)
Expo Hall Open (10 a.m. - 1:15 p.m.)
Member Council Meetings (10 a.m. – 11:30 a.m.)
Resolutions Committee Meeting (10:15 a.m. – 11:30 a.m.)
State League Steering Committee Meeting (11 a.m. – 12 p.m.)
Women In Municipal Government (WIMG) Membership Meeting (11 a.m. – 12 p.m.)
*Attendee Lunch in Expo Hall (12 p.m. – 1:15 p.m.)
Opening General Session (1:30 p.m. – 3 p.m.)
Expo Open (3 p.m. – 5 p.m.)
*Expo Hall Dessert Break (3 p.m. – 4 p.m.)
Conference Workshops (4 p.m. – 5:30 p.m.)
Constituency Group Meetings & Events (5:45 p.m. – 11 p.m.)
Young Municipal Leaders Networking Reception (9 p.m. – 11 p.m.)
NBC-LEO President's Hospitality Reception (9 p.m. – 11:45 p.m.)

FRIDAY, NOVEMBER 21, 2025

Registration Open (7 a.m. – 5:30 p.m.)
Ask NLC Open (8 a.m. – 6:30 p.m.)

Member Council Meetings (9:00 a.m. – 10:15 a.m.)
Expo Hall Open (10:00 a.m. – 1:30 p.m.)
Member Council Meetings (10:30 a.m. – 12 p.m.)
Conference Workshops (10:30 a.m. – 12 p.m.)
*Attendee Lunch in Expo Hall (12 p.m. – 1:30 p.m.)
Conference Workshops (2:15 p.m. – 3:45 p.m.)
Afternoon General Session (4 p.m. – 5:30 p.m.)
Constituency Group Events (5:30 p.m. – 11:30 p.m.)

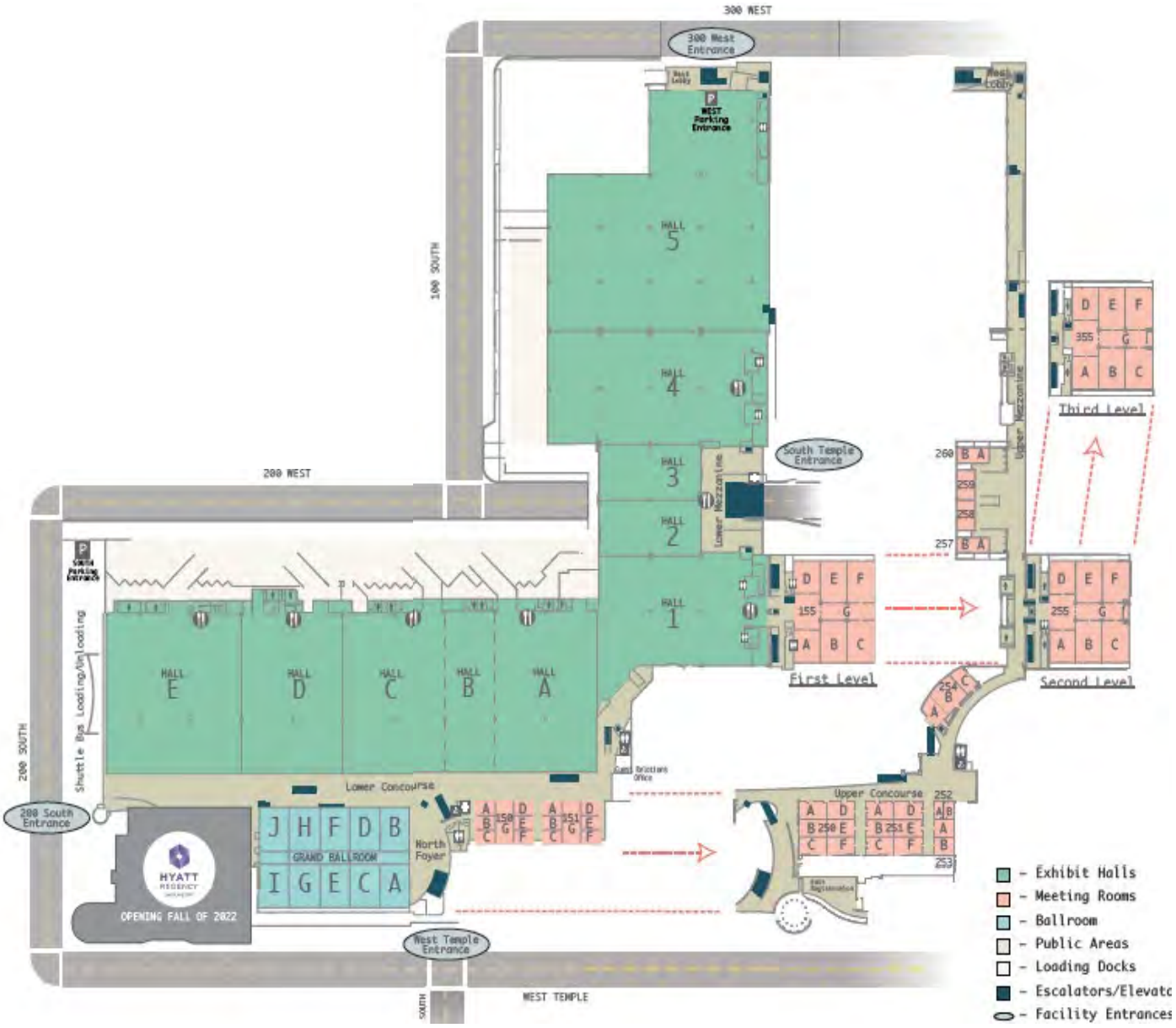
SATURDAY, NOVEMBER 22, 2025

Registration Open (7:30 a.m. – 1:30 p.m.)
Ask NLC Open (8:30 a.m. – 1:00 p.m.)
Conference Workshops (9 a.m. – 10:30 a.m.)
National Black Caucus Local Elected Officials (NBC-LEO) General Membership Meeting (10:45 a.m. – 12:15 p.m.)
Conference Workshops (10:45 a.m. – 12:15 p.m.)
*Closing Luncheon & General Session (12:30 p.m. – 2:15 p.m.)
NLC Annual Business Meeting (2:30 p.m. – 4:30 p.m.)
*Host City Closing Event (7 p.m. – 10 p.m.)

**Indicates functions which include conference-wide meals.*

View Workshop Descriptions View Full Conference Agenda

Calvin L. Rampton Salt Palace Convention Center Map



Hyatt Regency Floor Plan

HYATT REGENCY®

HYATT REGENCY SALT LAKE CITY
170 South West Temple
Salt Lake City, UT 84101 USA
T +1 888 591 1234
SLCRS-RFP@hyatt.com
www.hyattregencysaltlakecity.com

FLOOR PLAN

LEVEL 6

LEVEL 5

LEVEL 4

LEVEL 3

LEVEL 2