DOJ Report on AI in Criminal Justice: Key Takeaways

April 2025

Executive Summary

The rapid growth of artificial intelligence (AI) technologies presents both opportunities and challenges for the U.S. criminal justice system. As AI tools and applications are rapidly produced and deployed, criminal justice leaders find themselves trying to balance AI's transformative capabilities and its potentially harmful impacts.

The federal government's approach to AI governance underwent a significant shift in early 2025. On January 20, 2025, Executive Order 14148 revoked the previous Biden-Harris AI Executive Order 14110 of October 30, 2023 (Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence). This revocation was followed by Executive Order 14179 on January 23, 2025. Titled "Removing Barriers to American Leadership in Artificial Intelligence," the order established new U.S. policy priorities focused on enhancing America's AI dominance. This policy shift coincided with the announcement of the Stargate Project, a $500-billion private sector initiative led by SoftBank and OpenAI, with participation from major technology partners such as Oracle, NVIDIA, Microsoft, and Arm. The project aims to build extensive AI infrastructure across the country, beginning in Texas, with goals of creating hundreds of thousands of jobs and strengthening America's strategic AI capabilities. This massive private investment aligns with the new Trump administration's emphasis on reducing regulatory barriers and promoting U.S. leadership in AI development through market-driven approaches.

Amid these significant policy shifts and new private sector initiatives, the Department of Justice's December 2024 report on AI in criminal justice remains a valuable resource for understanding the fundamental opportunities and challenges that AI presents across the criminal justice system. The 77-page report's analysis of key applications addresses persistent operational and ethical considerations that exist independently of the regulatory framework. While the policy environment has evolved, the core technical, operational, and civil rights challenges identified in the report continue to warrant consideration.

This document summarizes the key findings and recommendations from the DOJ report to help stakeholders wrestle with critical questions about AI adoption in criminal justice. While policy frameworks continue to evolve, the technical, operational, and ethical considerations outlined here provide a foundation for informed dialogue and responsible implementation of AI systems in criminal justice settings.

This summary report was prepared with assistance from Claude, Anthropic's AI model (Claude 3.7 Sonnet). Claude provided suggestions on structure, clarity, and conciseness. All content decisions and final review were performed by humans.

Key Areas of Focus

The report examines four main applications of AI in criminal justice: (1) identification and surveillance; (2) forensic analysis; (3) predictive policing; and (4) risk assessment.

Identification and Surveillance[1]

This section examines how artificial intelligence could enable more effective biometric identification and surveillance across the criminal justice system and highlights critical concerns about accuracy, bias, privacy, and civil rights.

Key Applications

- **Automated Fingerprint Identification Systems:** Established biometric tool in use for over a century
- **Facial Recognition Technology:** Expanding since the 2010s, used for one-to-one identity verification and one-to-many database searches. Performance depends on image quality and environmental factors
- **Iris Scanning:** Review new biometric modality examining unique patterns in iris tissue in order to improve identification
- **Automated License Plate Recognition:** Widely used for real-time vehicle

identification and movement tracking. Operates through law enforcement and commercial networks, collecting billions of records annually

Key Benefits

- **Efficiency**
  - Makes analysis of large amounts of data feasible
  - Enables affordable identification of individuals and evidence
- **Accuracy**
  - Enables higher accuracy in identification through advanced pattern recognition and complex biometric data analysis
  - Allows faster, more reliable identification processes across investigation and supervision functions

Key Challenges

- **Performance and Bias**
  - Performance variations based on race, gender, age, and other characteristics
  - Need for comprehensive testing across different conditions and demographics
- **Privacy and Civil Rights Protection**
  - Data collection and retention policies must be clearly defined
  - Strict controls needed for database access and sharing
  - Special protections required for constitutionally protected activities
  - Commercial data use requires additional scrutiny

Recommendations

- Mandatory training for all system users

- Clear policies on authorized use cases
- Quality control procedures for system outputs
- Regular auditing and monitoring
- Transparent public reporting
- Continuous evaluation of system performance

These AI-enabled identification systems offer significant benefits for crime prevention and detection, but their implementation requires agencies to carefully balance the public safety advantages against privacy and civil rights concerns. Success depends on comprehensive policies, regular evaluation, and strong oversight mechanisms.

[1] Identification and Surveillance addresses both “police surveillance” and “prison-management tools.”

Forensic Analysis

The integration of artificial intelligence into forensic science represents a potential shift from subjective analysis toward more objective, reproducible approaches. While this transition offers potential improvements in accuracy and efficiency, it also introduces new challenges requiring careful validation and oversight.

Key Current Applications

- **Biometric Analysis**: Widely used for fingerprint, palm print, iris, and face comparison
- **DNA Analysis**: Probabilistic genotyping enables interpretation of complex samples containing mixed or degraded DNA
- **Narcotics Tracing**: The Drug Enforcement Administration employs machine learning to classify geographic origins of heroin and cocaine samples.
- **Digital Forensics**: AI assists in analyzing photos, videos, and communications, including detection of potential AI-generated content.

Key Emerging Applications

- **Pattern and Trace Evidence:** Used for analysis of toolmarks on bullets and cartridges, footwear impression comparison, glass fragment analysis, automotive paint analysis, and detection of ignitable liquids
- **Drug Evidence Analysis:** AI can enhance drug analysis to better identify and classify drug types
- **Forensic Medicine, Pathology, Anthropology, and Biology:** Used for age estimation from bruise analysis, sex and age determination, post-mortem identification, sperm cell detection in sexual assault evidence, and enhanced DNA mixture analysis
- **Crime Scene Investigation:** Used for automated categorization of crime scene photographs, enhancement of underwater imagery, and blood spatter pattern analysis

Key Benefits

- **Accuracy and Reproducibility**
  - Improves reproducibility and accuracy of forensic methods
  - Helps quantify the likelihood of matches and errors
- **Bias Mitigation**
  - Mitigates potential human biases and variation by standardizing processes
- **Efficiency**
  - Lowers costs associated with evidence examination
  - Reduces time required for complex forensic analysis processes

Key Challenges

- **Data Quality and Availability**
  - Requires large volumes of high-quality, representative data
  - Specialized equipment and samples often needed

- Data collection can be expensive and labor-intensive
- Privacy concerns with sensitive forensic data

- **Validation Requirements**
  - Must demonstrate methodological reproducibility and accuracy, both in principle and in particular cases
  - Requires rigorous testing for accuracy and potential biases
  - Need for continuous monitoring and revalidation, independent testing, and transparency
- **Explainability**
  - Current forensic AI models are generally interpretable so an expert is able to explain how specific inputs lead to particular outputs
  - More complex future models may present challenges for court testimony
  - Need to balance model complexity with explainability requirements
- **Human Oversight**
  - Essential for quality control and court admissibility
  - Risk of human biases affecting analysis
  - Need for specialized training to minimize bias
  - Importance of clear procedures and guidelines

Recommendations

- **Policy Framework:** Forensic science providers should establish clear AI policies, maintain human expert oversight and interpretation, and implement rigorous validation requirements and regular auditing of AI system use.
- **Procurement Requirements:** Procure only well-validated tools with demonstrated accuracy, require detailed vendor documentation, avoid restrictive licensing agreements, and ensure thorough bias evaluation
- **Training:** Provide comprehensive training on AI systems, share experiences between forensic providers, keep current with emerging technologies, and focus on bias

mitigation strategies

- **Datasets:** AI tools should be trained on large, high-quality, and representative datasets. Forensic science service providers should consider supplementing training data with jurisdiction-specific datasets when appropriate.

Predictive Policing

Predictive policing uses quantitative methods to identify likely crime locations, times, and perpetrators. At present, these tools inform resource allocation but do not prescribe specific interventions or evaluate their effectiveness. Predictive policing uses data analytics to forecast crime patterns but requires careful implementation to avoid perpetuating systemic biases.

Key Applications

- **Place-Based Models:** Place-based models focus on identifying potential crime hotspots by leveraging data to analyze where to focus limited policing resources.
- **Person-Based Models:** Person-based models attempt to identify people at elevated risk of offending or victimization. Many existing programs were halted by the end of the 2010s due to concerns regarding efficacy, bias, and civil rights.

Key Benefits

- **Effectiveness:** Promote public safety by informing focuses
- **Efficiency:** Enable more efficient use of limited resources
- **Transparency:** Enhance transparency and uniformity in decision-making processes

Key Challenges

- **Data Quality:** Historical crime data reflects existing biases and reporting patterns that often underrepresent certain crimes (e.g., domestic violence), are uncertain about basic data elements (time, location), and may reflect broader human and systemic biases.

- **Civil Rights Concerns:** Predictive policing may risk amplifying historical biases and feedback loops that can entrench discriminatory practices. The models may also have a disproportionate impact on vulnerable communities or erode public trust through increased surveillance.
- **Privacy Implications:** Enhanced scrutiny of identified individuals and data aggregation risks exposing sensitive information may require robust oversight mechanisms.
- **Interpreting and Explaining:** It is difficult for human to interpret and explain the internal workings of the most complex models, which makes accountability and oversight more complicated.

Recommendations

- **Community Partnership:** Engage stakeholders in goal setting, establish clear success metrics, maintain transparency in implementation, and seek out regular community consultation
- **Technical Implementation:** Practice careful data selection and validation, implement regular testing and independent auditing, focus on interpretable models over complex ones, and continuously monitor for unintended consequences
- **Operational Requirements:** Implement comprehensive staff training, clear governance policies, regular system reevaluation, and integration with non-police interventions

Risk Assessment

Risk assessment instruments use quantitative analysis to estimate outcomes in the criminal justice system, such as recidivism or failure to appear for trial. Benefits may include improved accuracy, enhanced efficiency for resource allocation and monitoring, increased transparency for risk calculation documentation and review, and greater equity that minimizes demographic disparities.

Key Applications

- **Pretrial Release:** Risk assessment tools help estimate whether defendants awaiting trial will fail to appear in court, commit new offenses, or pose public safety risks.
- **Sentencing:** Tools can help judges assess amenability to treatment, estimate recidivism likelihood when determining appropriate sentences within statutory ranges and guidelines.
- **Prison Classification:** Risk assessment informs decisions about facility placement, housing assignments, and program availability. While most current tools focus on post-release recidivism, some are specifically designed to predict misconduct during incarceration.
- **Probation and Parole Supervision:** Tools inform supervision levels, release conditions, and earned release eligibility for individuals not in custody or preparing for release.

Key Benefits

- **Accuracy:** Systematic evaluation of risk could be more accurate than subjective human judgments alone
- **Efficiency:** Improve system efficiency by directing costlier aspects of criminal justice monitoring or detention toward situations with greater potential benefit
- **Transparency:** Increase transparency through publicly accessible documentation, validation studies, and practitioner guidance
- **Equity:** Models can be designed and validated to minimize disparities across demographic groups and ensure similar estimates for individuals with similar characteristics

Key Challenges

- **Accuracy Limitations:** Even if these models augment and improve human judgment alone, these predictions are never certain. Just because someone is flagged as high risk doesn’t mean they will re-offend, and someone labeled low risk might still commit another crime. Life outcomes area also generally difficult to predict with accuracy.

- **Bias and Inequality:** Studies show performance differences can occur across demographic groups. Even when these tools seem equally accurate across different groups by one measurement, they might be biased when measured differently. Balancing accuracy with fairness is one of the biggest challenges when using risk assessment tools in criminal justice.
- **Data Quality Issues:** Data may be impacted by inconsistent collection and categorization practices, subjective inputs introducing potential bias, and limited evaluation of input consistency across practitioners.
- **Transparency Concerns:** Affected individuals may lack information about tool usage and functioning, and commercial protections may limit independent evaluation.
- **Validation Gaps:** Data may be affected by a lack of local validation and insufficient study of practitioner integration and community impacts.

Recommendations

- **Design and Implementation**
    - Document objectives, criteria, and alternatives before adoption
    - Engage affected communities throughout development
    - Carefully evaluate data quality and representation
    - Consider legal compliance regarding protected characteristics
    - Conduct phased deployment with independent evaluation
- **Ongoing Oversight**
    - Implement continuous monitoring of performance and bias
    - Maintain human decision-making authority
    - Provide notice and recourse to affected individuals
    - Conduct regular revalidation and make updates to account for changes
- **Training and Education**
    - Training should emphasize potential bias and disparities

  - Training programs should evolve with research advances and tool changes
  - All criminal justice stakeholders should understand tool functionality and limitations
- **Policy Requirements**
  - Issue clear guidance on appropriate use
  - Ensure public transparency about system design and testing
  - Enable meaningful access for independent research
  - Avoid tools with obscured functionality or inflexible thresholds
  - Provide decision-makers with detailed context beyond risk categories
- **Research**
  - More research is needed to understand risk assessment tools' performance and effects.

The report emphasizes that while risk assessment tools show promise for improving criminal justice outcomes, their implementation requires careful attention to accuracy, fairness, transparency, and ongoing evaluation. Success depends on robust safeguards and continuous monitoring to ensure that benefits outweigh potential harms.

Core Recommendations

Foundational Elements

- **Cost/Benefit Assessment**: Agencies should evaluate whether AI is the best solution, considering risks and benefits compared to alternatives.
- **Organizational Structure**: Agencies should define clear roles and responsibilities for AI oversight and implementation.
- **AI Inventory**: Maintaining a centralized record of AI systems, capabilities, and limitations is key in order to earn and maintain public trust. As feasible, agencies should aim for transparency by making their use cases public.

- **Workforce Training**: Agencies should ensure staff have technical and ethical AI expertise. Staff should understand operational considerations and potential impacts on civil rights and liberties as well as technical aspects.

Pre-Deployment Measures

- **Policy Development:** Agencies should establish clear guidelines on AI use, oversight, and compliance, including tool-specific policies.
- **Human Oversight:** AI should support, not replace, human decision-making—especially in high-stakes cases.
- **Community Transparency:** Agencies must disclose AI use, safeguards, and governance.
- **Data Integrity:** Agencies should ensure training data is relevant, unbiased, and representative.
- **Rigorous Testing:** AI systems must undergo operational testing and independent audits before full deployment.

Post-Deployment Measures

- **Ongoing Monitoring**: Regular audits for accuracy, bias, and real-world impact
- **New Use Evaluation**: Any expansion of AI applications must undergo reassessment.
- **Community Outreach**: Engage the public and relevant stakeholders regarding the intended uses of rights-or-safety impacting AI tools.

Cross-Cutting Implementation Considerations

- **Risk Management:** Higher-risk AI systems require stricter safeguards and oversight.
- **Technology-Specific Policies:** High-impact tools such as facial recognition need tailored regulations.
- **Documentation and Compliance:** Maintain records of system design, testing, regular

updates, and performance audits

- **Community Feedback:** Maintain open channels for public concerns and accountability

Conclusion

The Department of Justice's framework for AI governance in criminal justice emphasizes the need for agencies to strike a balance between proactive adoption of technological tools to pursue safety and justice and caution to protect fundamental rights.

The report says that successful implementation requires criminal justice agencies to establish robust organizational structures, to ensure public oversight and transparency, and develop appropriately trained workforces. Before deployment, agencies must implement detailed policies governing AI use, ensure robust human oversight, engage affected communities, and conduct thorough testing under real-world conditions. Following deployment, agencies must maintain rigorous monitoring protocols, regularly evaluate evolving uses, and sustain active community engagement.

The report acknowledges that AI applications in criminal justice will continue to evolve, particularly with emerging technologies like generative AI. It emphasizes that robust governance frameworks are essential for managing future developments while maximizing benefits and minimizing risks.

The recommendations outlined in this report, building upon previous federal guidance such as the NIST frameworks and OMB Memoranda on AI Use and Acquisition, provide a foundation for criminal justice agencies to develop and maintain effective AI governance programs. The report concludes that through careful attention to a set of core principles and practices, agencies can work to ensure that AI deployment enhances the fairness, effectiveness, and constitutional integrity of the criminal justice system.