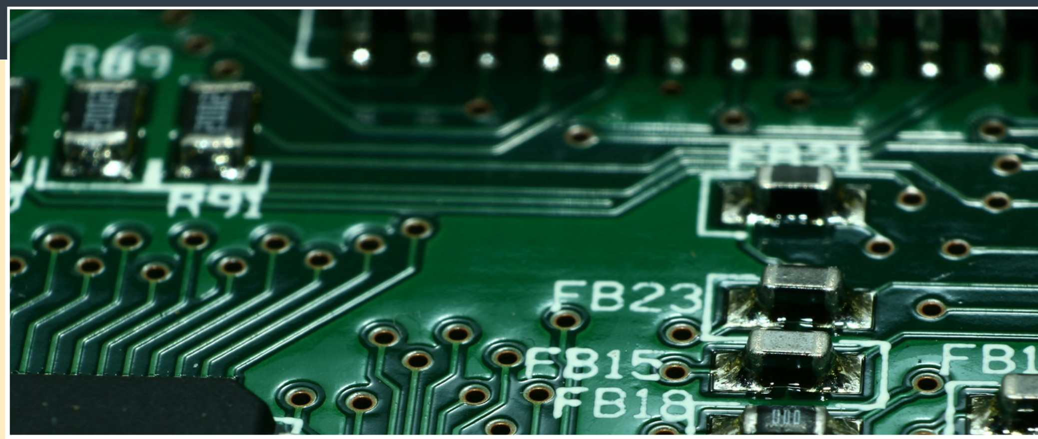
In-BRIEF

Forensic Technology
CENTER OF EXCELLENCE
A program of the National Institute of Justice

What FSSP Leaders Should Know About Artificial Intelligence and its Application to Forensic Science



"Technology is no longer the limiting factor here—it's the desire and the know how to integrate AI-enabled systems in a way that is fair and trustworthy across the board."

—Henry Swofford, PhD, Lead Scientist, Forensic Science Research Program, NIST

OBJECTIVES

- Introduce the concept of AI, its value to FSSPs, and current R&D efforts.
- Highlight the realities and considerations for implementing
- AI-enabled technologies to analyze physical forensic evidence.
- Identify what FSSP leaders should consider before developing and implementing AI-enabled technologies.

Introduction

Artificial intelligence (AI), which leverages computers to perform tasks that "enhance decision-making, problem-solving capabilities, and technology-driven innovativeness," is a rapidly growing area of interest for both industry and the federal research and development (R&D) community.[1] For forensic science service providers (FSSPs), AI-enabled technologies represent a significant opportunity to improve the way FSSPs identify, analyze, and reach conclusions on forensic physical evidence. Although most AI applications for forensic evidence currently remain in R&D phases, FSSP leaders must educate themselves and plan for future opportunities to invest in and leverage continuing AI advances.

The opportunities to improve forensic laboratory workflows and analyses with AI are significant, but AI and its potential implementation by FSSPs is multifaceted and complex. This brief was developed to help FSSP leadership understand the opportunities and realities of implementing future AI-enabled technologies in forensic science, highlight current research efforts, and help readers effectively plan for the inevitable future of AI.

Note to Readers: This brief focuses on the applications of AI toward analysis and interpretation of forensic evidence in physical disciplines. This document does NOT cover AI applications specifically for lab operations (e.g., resource allocation), use in analyzing digital evidence (e.g., extensive analysis of data obtained from computers, video feeds, internet activity), or use of generative AI in forensics (e.g., report writing).

Ten Points that FSSP Leaders Should Know About AI and its Application to Forensic Science

1. Applications to improve forensic analysis of physical evidence using AI are primarily in a conceptual or R&D phase.

As a continuously evolving discipline, AI encompasses an ever-growing number of techniques and applications. AI-enabled tools are now widely prevalent in many sectors, such as healthcare and banking, and integrated into many daily processes, such as using an internet search engine. Like the private sector, many federal agencies, including the Department of Justice, are expressing interest and investing in AI-enabled technologies that may be capable of optimizing resources and expanding capabilities across many industries. Organizations such as the National Artificial Intelligence Initiative and the National Science and Technology Council's Select Committee on Artificial Intelligence have laid the groundwork for awareness, strategic federal investment, and responsible and transparent implementation of AI-enabled technology. Facilitated by federal R&D investments, the criminal justice community has started to implement AI. Current applications center around use in law enforcement (e.g., predictive policing, facial recognition, automated license plate readers), criminal courts (e.g., risk assessments for pre-trial release and sentencing decisions), and corrections (e.g., prisoner communication monitoring, contraband detection, risk assessments for parole and early release decisions). Many digital evidence and data management companies have developed solutions for AI-enabled digital evidence recovery and analysis.[2] AI has remained an active area of R&D in the physical forensic science disciplines over the past 5 years (e.g., analysis of crime scene video, audio, and images), driven largely by National Institute of Justice investments. However, few forensics-focused research products have been commercialized or transitioned (or are ready for transition) into operational laboratories.

2. Current AI research focuses on machine learning (ML) techniques that may enhance forensic analysis capabilities or improve operational efficiencies.

AI applications include expert systems, ML systems, or hybrids of the two. ML and expert systems provide an output by feeding input data into an AI model (as shown in **Exhibit 1**). Expert systems use explicit rule-based reasoning or instructions developed by humans. ML systems use inferences derived automatically from data and find correlations and patterns in data that allow them to make future predictions and determinations about similar data.[3] Over the past 20 years, ML has had incredible success, and consequently current research in forensic applications of AI, like most applications of most federal AI investments, center around ML.[4] As such, this document will generally refer to AI-enabled tools for forensic applications as AI/ML tools.

Exhibit 1: Expert systems, supervised machine learning, deep learning, and unsupervised machine learning are examples of AI that have been applied in forensics.

AI tools provide an output by feeding input data into an AI model.

INPUT DATA → **PROCESS/MODEL** → **OUTPUT DATA**

AI applications relate input data, process/models, and output data in different ways. Forensic use cases of AI include the following.

AI Application		Description
Non-ML Based	Expert Systems	• Expert systems are rule-based systems where the process / model is explicitly defined.
ML-Based	Supervised Machine Learning	• Supervised learning models map inputs to outputs. These models are created by training on correct examples of input-output pairs, called "training data."
	Deep Learning	• Deep learning models use layers of artificial neurons, or nodes, to process information, recognize patterns, and develop predictions. It is often challenging to explain how a set of inputs resulted in the algorithmic output.
	Unsupervised Machine Learning	• Unsupervised learning finds underlying patterns and structures in data, rather than mapping inputs to outputs. It can be used to distill key information about data, such as finding clusters of related data points, or determining which features in the data are most important.

Most forensic research applications use an ML approach called supervised learning. Supervised learning can be used to classify inputs into certain categories or predict outputs based on inputs. In supervised ML, the AI/ML system is trained by finding the parameters that best map inputs to outputs using many labeled examples of input-output pairs, called "training data."[5] These training data must be carefully produced because biases, inaccuracies, and lack of diversity in training data will lead to poor accuracy in the final AI/ML tool (i.e., "garbage in, garbage out").

Supervised ML provides an opportunity for analysts to **expand the capabilities of detecting patterns and classifying and characterizing evidence.** Classification tools use defined attributes (e.g., measurable or observable characteristics) to place an item in one or more classes, such as a type of seized drug, type of body fluid, or biological sex. Regression tools, another ML subset, use defined attributes to estimate a continuous variable, such as postmortem interval, enabling the analyst to **predict or reconstruct information** based on a set of known examples (e.g., predicted age at death or age of a bruise). ML tools may help practitioners **make informed estimations even in circumstances where data may be incomplete** (e.g., single nucleotide polymorphism [SNP] profiles with high error rates).

Deep learning is a type of ML that uses artificial neural networks, a model architecture inspired by the way neurons are connected in the human brain. Deep learning techniques are often employed to tackle more challenging tasks like classifying two- or three-dimensional images, video, and audio files.[6] Convolutional neural networks (CNNs) are special types of deep neural network especially suited for images, video, or audio that looks for correlations between nearby pixels and audio samples. CNNs are frequently used for image-based analyses such as organ segmentation and facial recognition.[7] In forensic applications where examiners must draw conclusions through comparison, these tools can **help practitioners draw more objective analyses and conclusions.**

A smaller subset of ML tools use unsupervised learning. Unsupervised learning looks for relationships and patterns in input data without a corresponding set of output or labeled data.[8] For example, clustering techniques have been used to identify evolution of body parts through the decomposition process.[9] This category of ML tools may help uncover **patterns from large amounts of data that are too difficult to perceive by human interpretation. Unsupervised imputation methods can also help fill in gaps with incomplete data**; for example, imputation methods have been used to address missing osteometric data for forensic skeleton specimens.[10]

Constantly evolving technology and diverse perspectives make it challenging to universally categorize what "is" and "isn't" AI. A notable forensic example involves John Buckleton, one of the developers of STRmix probabilistic genotyping software, who published an open letter in a 2021 response to the Law Commission of Ontario's AI Case Study: Probabilistic Genotyping DNA Tools in Canadian Criminal Courts, which classified the tool as AI-based. Mr. Buckleton writes, "STRmix is not AI. AI has no official definition but generally is considered to emulate some aspects of human intelligence. Most often this involves some aspect of learning, sensing the environment, or judgement. STRmix does none of these, is neither based in machine learning nor other heuristic approaches, and makes no 'decisions.'"

Regardless of this distinction, this probabilistic genotyping tool is part of an "automated system," and FSSP leaders should understand potential implications for bias and other considerations prior to adopting.

3. Some FSSPs are currently adopting automated tool-based technology that does not fit into the machine-learning technology category but raises similar implementation challenges.

The forensic community has implemented technologies that use algorithms, statistical models, and other computational tools to help identify patterns and enable decision-making. However, leadership should understand that not all forensic tools using algorithms and computational models fit under the ML umbrella. For example, FSSPs have implemented commercially available forensic algorithms for probabilistic genotyping and latent print analysis as tools that may help "assess whether or not evidence collected in a criminal investigation may have originated from an individual."[11] Although these tools themselves do not fit under the ML umbrella, ML techniques may be used in conjunction with existing latent print or probabilistic genotyping tools. The White House Office of Science and Technology Policy's Blueprint for an AI Bill of Rights refers to both examples of tools (ML and non-ML based) as "automated systems," or "any system, software, or process that uses computation as whole or part of a system to determine outcomes, make or aid decisions, inform policy implementation collect data or observations, or otherwise interact with individuals and/or communities."[12]

Automated tools, whether ML-based or non-ML based, are subject to biases, depend significantly on the parameters established and the quality of data used during development, and require transparency and improved policies for testing, performance, and use. FSSP leaders considering ML technology implementation should also consider the implications of automating points of the decisionmaking process before identifying whether ML or non-ML-based tools address laboratory needs. FSSPs should use the automated tool that approaches their operational needs in the simplest way.[12]

4. The development and implementation of AI/ML-enabled technology is a paradigm shift from traditional forensic analysis methods—and FSSP leadership should clearly define a tool's purpose and fit into existing workflows.

The continuous improvement of forensic science traditionally relies on the scientific method, a dedicated process that includes developing, testing, and refining a hypothesis. Although AI/ML helps reveal patterns in data, it does not ultimately reveal why these data may be related in a way that people can understand. Laboratory leadership and technical staff monitoring future AI applications must understand that this "paradigm shift" from traditional methods has implications for applications and effective use cases of this technology.

Although AI/ML-enabled technologies hold high-level potential to optimize the provision of forensic science, true value depends on how the tool will be used and where it is implemented within the workflow. Media coverage driving the AI hype paints a lofty envisioned future of AI/ML automating tasks without the need for human intervention. Realistic implementation of AI/ML tools, however, exists on a spectrum. Swofford et al. described a similar spectrum for implementation of forensic algorithms, an adjacent technology area, which the Forensic Technology Center of Excellence has modified for specific AI/ML technologies (see Exhibit 2).[13] FSSP leadership looking to implement AI should consider their objectives in adopting the technology: Is the intention to increase efficiency of analysis, to provide a "safety net" to support conclusions made with traditional methods, or to enhance capabilities beyond what was traditionally possible? Although the up-front consideration on workflow integration should be led by FSSP leadership and technical experts, researchers and companies offering AI/ML-based technologies can help FSSPs understand appropriate and realistic applications for their technology; Foster + Freeman, for example, offers "AI-Assist" software that is built into their Amino Acid Rapid Imager (AARI®) system for fingerprint examination. AI-Assist was created to speed up the process of the human examiner, not to directly identify a fingerprint match.

"Traditional genetic practitioners consider computer science the same as statistics. However, traditional statistics is an old-fashioned way of using AI. Latest developments have little to do with traditional AI. People in the forensic community should consider the distinction"

—Jianye Ge, Independent Consultant, Forensic Biology Discipline

Implementing technologies with the intention of more ML involvement in current workflows will naturally take more forensic community buy-in, up-front planning, testing and evaluation, and governance. Realistically, the "sweet spot" for most forensic AI applications may land between Levels 1–3, where the examiner plays a key role in the evaluation of the evidence (Exhibit 2). In levels 4-5, where examiners shift their role from making conclusions to overseeing the AI/ML tool make the conclusion, the effort needed to shift workflows, policies, and build in effective quality assurance measures may be prohibitively high.

Exhibit 2: Potential Spectrum of AI/ML Involvement in Forensic Science Applications.

Level of AI/ML Involvement	Description	Example Use Cases in Forensic Science Analyses
0: No AI	Humans rely on collective knowledge to operate tasks and make decisions without any intervention of AI/ML-enabled technology, including the formation of an expert opinion.	Examiners use non-AI/ML-based methods to arrive at conclusions.
1: AI Assistance	AI/ML tools are incorporated as an optional workflow step after an examiner has made an expert opinion.	An examiner may use traditional methods to determine whether two prints may have originated from the same source, then leverage an AI/ML-based tool to reinforce or reconsider findings before issuing a report. This use case would be similar to consulting an opinion of another examiner.
2: AI as Quality Control	AI/ML tools are incorporated as a mandatory step after an examiner has made an expert opinion.	An examiner must use AI/ML-enabled tools for latent print analysis to ensure that results of the analysis meet quality assurance standards.
3: AI-Informed Evaluation	AI/ML is used as an optional supplemental factor to inform the expert opinion (i.e., before the opinion is made).	Examiners may use AI/ML-enabled tools to help make them reach conclusions more effectively (e.g., to use these as screening tools which are confirmed through further analysis).
4: AI-Dominated Evaluation	AI/ML is used as the primary basis for the evaluation of evidence (i.e., mandatory tool driving the opinion made), and the examiner oversees the application of the tool.	Examiners must use AI/ML-enabled tools to reach a conclusion, especially in circumstances where human-based interpretation is difficult or not feasible.
5: Full AI Implementation	AI is used to develop a conclusion without human oversight.	AI/ML-enabled tools reach conclusions without examiner oversight or intervention.

5. Most forensic applications for AI/ML focus on enabling more objective analyses of impression and pattern evidence or classification and prediction of features based on qualitative or quantitative data.

Researchers across several forensic disciplines are looking to incorporate ML, deep learning, and other AI tools to address key challenges in forensic science. The following section provides an overview of emerging technology and areas of specific impact.

Impression and Pattern Applications

AI/ML presents an opportunity to help analysts arrive at a decision (i.e., Levels 3 and 4) as to whether the source of an impression or a pattern image at the crime scene (e.g., latent fingerprint or a toolmark) matches an impression or pattern from a specific suspect or weapon and the likelihood or probability of that match being true. Impression and pattern evidence varies but makes up the largest set of potential AI/ML applications; sub-disciplines include analysis of firearm casings and bullet marks, toolmarks, latent fingerprints, questioned documents, and blood spatter.

Although current methods rely on the subjective approach of visual analysis and interpretation, ML can be used to help quantify the similarity between two items and the "frequency with which a given degree of similarity between two items can be expected when the items have a common source and when they do not."[14] Researchers have explored the use of classification algorithms for assessing similarity of lands engraved in fired bullets,[15] similarity of footwear impressions based on a database of outsoles impressions with diverse degrees of wear,[16] and similarity of writing marks from graphite pencils based on laser-induced breakdown spectroscopy results.[17] Facial recognition, an adjacent AI/ML application in the criminal justice community, is perhaps one of the most widely known implemented applications of AI/ML for pattern (image) analysis and identification.

AI/ML is of significant interest to many researchers in the latent print field, with several researchers investigating the ability to assess similarity, address challenges of photometric and geometric distortion, and explore ridge reconstruction.[18] AI/ML methods for improving fingerprint identification have been tested in biometrics and security purposes, although the use case scenarios are quite different from latent fingerprint analysis in forensics. Most commercial or operational AI/ML-based forensic tools are related to latent print analysis. For example, Automated Fingerprint Identification Systems (AFIS) are widely used around the world and are built on various algorithmic approaches, including ML methods. AI/ML implementation in AFIS has driven advancements such as image enhancement, feature extraction, indexing, and matching. Latent Quality Metric (LQMetric), developed by Nobilis, is an ML-enabled tool that characterizes the quality of ridge detail in latent fingerprint images and has been implemented in the FBI's Universal Latent Workstation software.[19]

Beyond pattern matching, researchers have explored the use of AI/ML techniques to help them classify, automate, and ultimately improve qualitative assessments. Recent research efforts include use of deep CNNs to classify descriptors of shoeprints left at crime scenes (to automate coding for a database),[20] classify blood patterns,[21] and ultimately estimate the angle of impact from blood spatter.[22]

DNA and Biology Applications

AI/ML represents an opportunity to help support analyst interpretation of challenging DNA analyses. Multiple researchers have investigated the use of classification techniques to improve existing probabilistic genotyping algorithms, using a dataset of known mixture samples to identify the number of contributors in unknown samples.[23,24] Currently, Niche Vision's PACE software is the only known DNA-related software on the market that integrates ML in DNA analysis, but it is currently being incorporated into emerging analysis tools for new approaches to probabilistic genotyping, including unique molecular identifier "barcodes" that may provide more insights into polymerase chain reaction and sequencing errors.[25] AI/ML may help analysts distinguish signal from noise to enable complex low-level interpretation.[26]

AI/ML may help analysts develop informed predictions of sample source and other important traits—for example, researchers have employed classifier tools to predict externally visible characteristics from DNA datasets, including eye, skin, and hair color,[27] and to identify biological source of a sample based on microbial signature. AI/ML is also being used for age prediction through assessing patterns in DNA methylation especially in samples that are degraded or low quantity.

The emerging field of investigative genetic genealogy is an area where AI/ML techniques may be especially helpful because the process involves generating and interpreting SNP data and identifying genealogical relationships based on the SNP data. For example, researchers are estimating genealogical relationships in circumstances where the resulting SNP profile contains significant errors (e.g., missing persons cases).[28]

Anthropology Applications

AI/ML techniques can help forensic anthropologists incorporate several measurements and inputs to infer important traits, especially from morphological and morphometric characteristics of bone. Many researchers have developed tools leveraging classification techniques and measurements from human skull measurements[29] and tooth morphology[30] to estimate an individual's ancestry and sex determination.[31] AI/ML methods have been implemented in age at death estimation models of dental measurements in subadults to address challenges in missing data[32] and measurements of the pubic symphysis.[33] Deep learning tools have enabled researchers to estimate sex and age at death using actual 2D and 3D images like CT scans, with the ability to measure and recognize landmarks and other relevant image elements.[31,34] AI/ML can also play a role in identifying decedents when other traditional methods fail, including postmortem iris recognition[35] and facial recognition for individuals who are found in a state of postmortem decomposition.[36]

Pathology Applications

AI/ML may help pathologists interpret data using advanced imaging technologies. CNNs and other deep learning tools can help distinguish fatal head injury (e.g., subarachnoid hemorrhage) in postmortem computed tomography[37] and age bruises visible or invisible to the naked eye.[38] Researchers have employed supervised ML techniques for segmentation of body parts[39] and organ images and have also considered unsupervised learning to cluster similar images of decomposing bodies to help assist manual annotation of features. Beyond image analysis, pathology applications include analysis of RNA, mRNA, miRNA, and proteomic data to suggest certain pathologies, such as acute myocardial ischemia, and to estimate wound age[40] and postmortem interval[41] from microbial markers.

Crime Scene Applications

Researchers are currently exploring the use of CNNs and other deep learning algorithms to improve visualization and detection of evidence at the scene; for example, CNNs may help remove noise and enhance sharpness of underwater images, helping investigators find and capture evidence in underwater crime scenes.[42] CNNs may also help researchers classify weapons based on audio source of a muzzle blast or help build a content-based image retrieval system capable of classifying crime scene images such as weapons or illicit drugs.[43] These technologies may enable field analysis and interpretation of materials at the scene by enabling classification of complex spectra generated by field instrumentation; for example, researchers developed models to accurately classify ignitable liquids based on handheld Raman spectrometer data.[44]

Toxicology and Seized Drugs Applications

AI/ML-enabled technologies provide an opportunity to improve identification of compounds found in seized drugs and toxicological samples and ultimately keep up with the "arms race" of identifying novel psychoactive substances and other new drugs hitting the market. For example, researchers have employed CNNs to classify spectra from portable analysis instrumentation, such as Raman spectrometers, to presumptively detect and identify fentanyl-related compounds.[45] Beyond screening, ML techniques can help classify and ultimately identify novel fentanyl analogs from mass spectrometry[46] or gas chromatography-mass spectrometry.[47] ML also provides an opportunity for the toxicology community to predict drug toxicity based on symptoms presented by a patient as a screening mechanism.[48]

Trace Evidence Applications

ML algorithms can help trace evidence analysts classify and identify materials. Researchers used deep learning to classify manufacturer and assembly plants for clear coat automotive paint formulations, which are challenging to identify via Fourier Transform Infrared Spectroscopy spectra or automotive paint databases alone.[49] Similarly, researchers have used ML for nondestructive identification of heavy mineral oil based on Fourier transform Raman spectroscopy.[50] Beyond spectra, AI tools may be able to classify evidence such as types of glass fragments left at a scene.

6. AI/ML-enabled tools are often black boxes that have significant implications for reliability and court admissibility.

AI/ML tools are often based on black box methods with complex inner workings that may make it extremely challenging to understand how the tool produced a particular result. These methods infer relationships and rules based on patterns and correlations in data, rather than hypothesizing, testing, and refining explicit causal relationships as traditional scientific methods do. Results may sometimes be based on spurious correlations. Additionally, limitations in the data used to develop AI/ML tools (e.g., data not representative of an entire population) can introduce systematic biases in results. These limitations can call into question the reliability of these methods.

As more companies and developers begin to offer AI-enabled software products, hold them accountable. Consider including language in your procurement contracts to facilitate access to this source code when needed, and let your purchasing power push emerging companies into sharing this information.

This challenge is further complicated by vendor or developer unwillingness to share the source code with the court systems, deeming it a "trade secret."[51] For example, the 2021 *New Jersey State vs. Pickett* case denied a motion to request the source code of TrueAllele, a probabilistic genotyping software, on the basis that this source code constituted trade secrets.[52] In light of these challenges, establishing greater transparency in AI/ML-based tools is a widely discussed topic across policymakers.[a]

Use of information gathered with the help of AI/ML tools has been an active area of discussion with respect to civil liberties, civil rights, and the Federal Rules of Evidence (FRE).[53] Potential issues may arise around the First, Fourth, Fifth, and Sixth Amendments and the FRE. The Fifth Amendment's Due Process Clause prohibits the federal government from depriving anyone of life, liberty, or property without the due process of law. Accordingly, the inexplicability of AI/ML-generated evidence may be in opposition to the Amendment. However, it is likely that the court would not recognize this view because it would be difficult to establish that evidence must be *fully* explainable, whereas the technology's processes, methodology, data, and assumptions can *mostly* be explained and understood. Additionally, it is likely that ML output would be presented in the form of expert testimony, which would allow the defendant to cross-examine the expert on the technology's processes and capabilities. Moreover, the Sixth Amendment's Confrontation clause would require an in-person testimony that is subject to cross-examination. Other forms of testimony, such as requirements for expert witnesses to testify on drug analysis evidence, could be leveraged as a framework for how ML experts would testify in court.[53] Experts generally agree that these legal requirements would not pose a universal barrier to AI/ML evidence admissibility but instead could limit how it may be introduced. Evidence gathered from AI/ML-enabled and algorithmic tools are likely able to meet requirements for admissible expert testimony[51]; however, FSSP leadership should consider these admissibility requirements the minimum standard and not equate this to confirmation that AI tools are robust, accurate, and effective. The forensic community should discuss these challenges and consider scenarios that may impact use of AI/ML (e.g., considering a threshold that determines whether evidence generated via AI/ML technologies would be sufficient evidence for probable cause for arrest).

[a] Select examples of policymaker activity around AI transparency include the proposed Justice in Forensic Algorithms Act of 2021, which would establish a federal standard for testing computational forensic software and prohibit trade secret privileges to prevent defense access to evidence, including source code, in criminal proceedings. The Bipartisan Framework for U.S. AI Act calls for establishing an oversight body to regulate vendors operating in "high-risk" areas (e.g., facial recognition), including data governance.

7. Independent technology validation is a critical barrier to AI (and forensic algorithm) implementation.

Like other emerging technologies, AI/ML-enabled tools must be internally validated by FSSPs to assess the ability of the tool to perform in its intended use case. Few resources exist to help FSSPs design and execute these studies, which may require robust statistical analyses. Within the spheres of law enforcement–based AI tools and automated systems for forensic applications, policy experts have noted vendor reluctance to expose the inner workings of their algorithms or provide source code for review or testing. Many vendors elect to conduct validation studies of their own technologies, but FSSPs should consider these results in light of the vendor's potential motivations to present their product positively. Lack of validation studies may drive a lack of community understanding on best practices for implementing the technology.

Developing technical capacity for validation is a shared responsibility between the forensic researchers developing the tools and the FSSP evaluating and implementing them. For AI/ML tool developers, it is important to robustly quantify and publish the error rates of their tools and to perform these evaluations using test data that accurately represent real application data. For users of AI/ML tools, it is important to carefully examine the reliability of new tools before adoption and to clearly document and communicate when particular AI/ML tools have been used to reach specific determinations or decisions. In examining the reliability of new tools, it is best practice to consult with third-party experts. Together, developers and forensic community end users should consider the ever-evolving nature of AI/ML-based tools and collectively determine thresholds of what is considered a "validated tool," with the intention of providing guidance that defines when a tool lies within or outside of being "validated."

8. AI/ML-enabled technologies are subject to bias and other limitations, which require careful consideration of how and when these tools are used.

FSSPs need to consider policies for responsible use of AI/ML in light of their technical limitations, which include the proclivities for biases. Applications for facial recognition technology, which typically use AI/ML-based techniques, raise many important factors for FSSPs to consider when implementing these tools. Many facial recognition platforms are consistently and significantly identifying female individuals and persons of color less accurately.[54] These inaccurate results may lead to wrongful convictions and other harmful outcomes. Without proper understanding of responsible use and limitations of these tools, the forensic community may interpret results as "true," leading to inaccurate or unjust outcomes and community mistrust. FSSP leaders should work across the justice sector to consider and implement these tools in a fair and transparent way. NIST, for example, has developed a Trustworthy and Responsible Artificial Intelligence Resource Center to help educate and establish a roadmap toward responsible use of AI tools.

9. Successful implementation relies on buy-in from both leadership and the technical staff operating these tools.

Implementing AI/ML technologies into forensic techniques requires buy-in from all stakeholders, from both leadership and technology end users and individuals and communities that may be potentially harmed by the technologies. For the former, factors to consider include commercial viability and return on investment. For the latter, the main factor is considering what practitioners might need to feel comfortable and confident in the new technology.

"Though not technical, the greatest challenge is the mentality of the forensics community, in accepting AI in their workflows, as the potential for many people's work being replaced exists"

—Jianye Ge, Independent Consultant, Forensic Biology Discipline

Some practitioners may be wary of fully adopting AI/ML techniques to analyze evidence because of concerns about explaining the evidence in court. Although there are myriad reasons for this opposition, a main source of skepticism is the perception that AI/ML can only be implemented with an "all or nothing" approach (either human or the algorithm).[13] It is also important to consider the context in which those techniques will be applied; the outputs of those techniques will lead to sensitive decisions that may have significant impacts on the life of individuals. This pressure further adds to the reluctance to move away from traditional methods. There are many other anecdotal reasons for the skepticism, including perceptions of the inability of algorithms to incorporate qualitative data or consider individual circumstances. Those doubts can be addressed by creating a solid foundation for implementation with plans for education, training, protocols, validation, verification, and ongoing monitoring systems.

10. What now? To prepare for the AI wave, FSSP leadership should take initial steps toward AI/ML familiarity.

AI has many potential applications within forensic science. To prepare for implementing these technologies, FSSP leadership should take the following steps:

Stay on top of emerging AI/ML tools and techniques, monitoring and leveraging the "early adopters." Leadership should monitor AI/ML advancements in forensic applications by monitoring grant funding mechanisms, conference topics, and journal articles. Interlaboratory peer discussion about AI/ML implementation, in the form of working groups or round tables, can help the community collectively create a path forward for responsible implementation. Champions of technology implementation in adjacent criminal justice applications (e.g., digital evidence and investigation technologies) may serve as helpful resources as forensic-specific technologies reach the market.

Develop or acquire training resources that enable practitioners to effectively vet, validate, and use the AI/ML tool for its intended use case. A properly validated and admissible system is not enough to guarantee success. Before new AI/ML techniques are fully adopted, it is essential that lab leaders and researchers gain the foundational skills to effectively use the technology. Before implementing AI in casework, practitioners should understand the technical background of the technology, what to interpret (and not interpret) from results on an AI/ML-enabled tool, potential biases that could influence results from the tools, and how to test and evaluate emerging technologies.

Consider partnerships with AI/ML researchers to help test, evaluate, and validate technologies that are created for forensic use cases. Forensic AI/ML application is primarily in the R&D phase; to develop tools that align with practitioner needs, researchers must develop collaborative partnerships with forensic laboratories. These partnerships drive discussions that help researchers understand needs, enable data sharing (which may help train models), and help forensic laboratories build technical capacity and understanding of the tools.

Leverage a multidisciplinary team to plan for use cases and consider admissibility and reliability. Responsible implementation of AI/ML in forensic laboratories requires a multifaceted approach. Forensic laboratories should integrate the perspectives of legal (e.g., district attorneys, defense attorneys, judges), ethics, computer science, and information technology database experts to obtain a comprehensive understanding of how this technology will not only impact workflows but also uphold ethical standards embedded within society and judicial systems.

Pilot and familiarize the lab with AI/ML through other "lower-stakes" use cases. AI-enabled tools have the potential to not only improve forensic science analyses but also augment administrative tasks such/ML as budgeting, writing, and allocating resources (i.e., adjacent applications that are not the focus of this in-brief). AI/ML may also help report drafting and other writing tasks. By implementing currently available AI tools in non-casework applications, the team can familiarize themselves with these tools, understand benefits and limitations, and approach forensic applications in a better-informed way.

References

1 Bahoo, S., Cucculelli, M., & Qamar, D. (2023). Artificial intelligence and corporate innovation: A review and research agenda. *Technological Forecasting and Social Change*, 188. https://doi.org/10.1016/j.techfore.2022.122264.

2. Criminal Justice Testing and Evaluation Consortium. (2020). *Artificial intelligence in the Criminal Justice System*. https://cjtec.org/artificial-intelligence-in-the-criminal-justice-system/

3. Executive Office of the President of the United States. (2019). *The National Artificial Intelligence Research and Development Strategic Plan: 2019 Update*. https://www.nitrd.gov/pubs/National-AI-RD-Strategy-2019.pdf

4. Criminal Justice Testing and Evaluation Consortium. (n.d.). *Demystifying artificial intelligence, its applications, and potential risks*. https://cjtec.org/files/64bfb233b0d66

5. Zhou, A., & Bohan, O. B. (n.d.). *Machine learning* (ML). https://courses.cs.washington.edu/courses/cse160/18sp/sections/09/MachineLearning.pdf

6. Iterators. (2021). *Machine learning vs. Deep learning: The ultimate comparison*. https://www.iteratorshq.com/blog/machine-learning-vs-deep-learning-the-ultimate-comparison/

7. IBM. (n.d.). *What are convolutional neural networks?* https://www.ibm.com/topics/convolutional-neural-networks#:~:text=the%20intended%20object.-,Convolutional%20layer,matrix%20of%20pixels%20in%203D

8. Curtis, S. (2023). *How to implement enterprise resource planning (ERP)*. https://www.ibm.com/blog/

9. Mockus, A., & Steadman, D. W. (2020). ICPUTRD: *Image cloud platform for use in tagging and research on decomposition (report no. 255312)*. . National Criminal Justice Reference Service, Office of Justice Programs. https://www.ojp.gov/pdffiles1/nij/grants/255312.pdf

10. Pang, J., & Liu, X. (2023, Aug). Evaluation of missing data imputation methods for human osteometric measurements. *Am J Biol Anthropol*, 181(4), 666-676. https://doi.org/10.1002/ajpa.24787. PMID: 37259623.

11. United States Government Accountability Office. (2021). Algorithms strengthen forensic analysis, but several factors can affect outcomes.

12. Office of Science and Technology Policy, The White House. (2023). *Blueprint for an AI Bill of Rights: Making Automated Systems Work for the American People*. Retrieved January 12, 2023, from https://www.whitehouse.gov/ostp/ai-bill-of-rights/

13. Swofford, H., & Champod, C. (2021). Implementation of algorithms in pattern & impression evidence: A responsible and practical roadmap. *Forensic Sci Int Synerg*, 3, 100142. https://doi.org/10.1016/j.fsisyn.2021.100142. PMCID: PMC7933265. PMID: 33718855.

14. Carriquiry, A., Hofmann, H., Tai, X. H., & VanderPlas, S. (2019). Machine learning in forensic applications. *Significance*, 16(2), 29-35. https://doi.org/10.1111/j.1740-9713.2019.01252.x.

15. Hare, E., Hofmann, H., & Carriquiry, A. (2017). Algorithmic approaches to match degraded land impressions. *Law, Probability and Risk*, 16(4), 203-221. https://doi.org/10.1093/lpr/mgx018.

16. Park, S., & Carriquiry, A. (2022, Feb). The effect of image descriptors on the performance of classifiers of footwear outsole image pairs. *Forensic Sci Int*, 331, 111126. https://doi.org/10.1016/j.forsciint.2021.111126. PMID: 34922283.

17. Chen, Y., Liu, Y., Han, B., Yu, W., & Wan, E. (2022). Identification of writing marks from pencil lead through machine learning based on laser-induced breakdown spectroscopy. *Optik*, 259. https://doi.org/10.1016/j.ijleo.2022.169008.

18. National Institute of Justice. (2022). *An analysis on adversarial machine learning: Methods and applications*. https://nij.ojp.gov/library/publications/analysis-adversarial-machine-learning-methods-and-applications

19. Kalka, N. D., Beachler, M., & Hicklin, R. A. (2020). *LQMetric: A latent fingerprint quality metric for predicting AFIS performance and assessing the value of latent fingerprints*. Noblis. https://noblis.org/wp-content/uploads/2023/05/LQMetric_JFI-Final-2020-10-24.pdf

20. Budka, M., Ashraf, A. W. U., Bennett, M., Neville, S., & Mackrill, A. (2021). Deep multilabel CNN for forensic footwear impression descriptor identification. *Applied Soft Computing*, 109. https://doi.org/10.1016/j.asoc.2021.107496.

21. Liu, Y., Attinger, D., & De Brabanter, K. (2020, May). Automatic classification of bloodstain patterns caused by gunshot and blunt impact at various distances. *J Forensic Sci*, 65(3), 729-743. https://doi.org/10.1111/1556-4029.14262. PMID: 31944296.

22. Niketha, R. (2021). Development of an artificial intelligence method for the analysis of bloodstain patterns.

23. Marciano, M. (2016). *A hybrid machine learning approach (MLA) for DNA mixture interpretation*. https://www.aafs.org/sites/default/files/media/documents/AAFS-2016-B99.pdf

24. Alotaibi, H., Alsolami, F., & Mehmood, R. (2021). DNA profiling: An investigation of six machine learning algorithms for estimating the number of contributors in DNA mixtures. *International Journal of Advanced Computer Science and Applications*, 12(11).

25. Woerner, A. E., Mandape, S., King, J. L., Muenzler, M., Crysup, B., & Budowle, B. (2021, Mar). Reducing noise and stutter in short tandem repeat loci with unique molecular identifiers. *Forensic Sci Int Genet*, 51, 102459. https://doi.org/10.1016/j.fsigen.2020.102459. PMID: 33429137.

26. Marciano, M. A., Williamson, V. R., & Adelman, J. D. (2018, Jul). A hybrid approach to increase the informedness of CE-based data using locus-specific thresholding and machine learning. *Forensic Sci Int Genet*, 35, 26-37. https://doi.org/10.1016/j.fsigen.2018.03.017. PMID: 29627762.

27. Katsara, M. A., Branicki, W., Walsh, S., Kayser, M., Nothnagel, M., & Consortium, V. (2021, Jul). Evaluation of supervised machine-learning methods for predicting appearance traits from DNA. *Forensic Sci Int Genet*, 53, 102507. https://doi.org/10.1016/j.fsigen.2021.102507. PMID: 33831816.

28. Huang, M., Liu, M., Li, H., King, J., Smuts, A., Budowle, B., & Ge, J. (2022). A machine learning approach for missing persons cases with high genotyping errors. *Front Genet*, 13, 971242. https://doi.org/10.3389/fgene.2022.971242. PMCID: PMC9573995. PMID: 36263419.

29. Navega, D., Coelho, C., Vicente, R., Ferreira, M. T., Wasterlain, S., & Cunha, E. (2015, Sep). AncesTrees: ancestry estimation with randomized decision trees. *Int J Legal Med*, 129(5), 1145-1153. https://doi.org/10.1007/s00414-014-1050-9. PMID: 25053239.

30. Scott, G. R., Pilloud, M., Navega, D., Coelho, J., Cunha, E., & Irish, J. (2018). rASUDAS: A new web-based application for estimating ancestry from tooth morphology. *Forensic Anthropology*, 1(1), 18-31. https://doi.org/10.5744/fa.2018.0003.

31. Thurzo, A., Kosnacova, H. S., Kurilova, V., Kosmel, S., Benus, R., Moravansky, N., Kovac, P., Kuracinova, K. M., Palkovic, M., & Varga, I. (2021, Nov 12). Use of advanced artificial intelligence in forensic medicine, forensic anthropology and clinical anatomy. *Healthcare (Basel)*, 9(11). https://doi.org/10.3390/healthcare9111545. PMCID: PMC8619074. PMID: 34828590.

32. Herrmann, N. P., Hefner, J. T., Spence, J. E., & Kamnikar, K. (2023). *Investigation of subadult dental age-at death estimation using transition analysis and machine learning methods*. https://www.ojp.gov/pdffiles1/nij/grants/306558.pdf

33. Gámez-Granados, J. C., Irurita, J., Pérez, R., González, A., Damas, S., Alemán, I., & Cordón, O. (2022). Automating the decision making process of Todd's age estimation method from the pubic symphysis with explainable machine learning. *Information Sciences*, 612, 514-535. https://doi.org/10.1016/j.ins.2022.08.110.

34. Cao, Y., Ma, Y., Yang, X., Xiong, J., Wang, Y., Zhang, J., Qin, Z., Chen, Y., Vieira, D. N., Chen, F., Zhang, J., & Huang, P. (2022). Use of deep learning in forensic sex estimation of virtual pelvic models from the Han population. *Forensic Sci Res*, 7(3), 540-549. https://doi.org/10.1080/20961790.2021.2024369. PMCID: PMC9639534. PMID: 36353321.

35. Czaika, A. (2023). *Software tool and methodology for enhancement of unidentified decedent systems with postmortem automatic iris recognition*. https://www.ojp.gov/library/publications/software-tool-and-methodology-enhancement-unidentified-decedent-systems

36. US. Department of Energy, Office of Scientific and Technical Information. (2019). *Effects of postmortem decomposition on face recognition*. https://www.osti.gov/biblio/1559672

37. Garland, J., Hu, M., Kesha, K., Glenn, C., Morrow, P., Stables, S., Ondruschka, B., & Tse, R. (2021, Mar). Identifying gross post-mortem organ images using a pre-trained convolutional neural network. *J Forensic Sci*, 66(2), 630-635. https://doi.org/10.1111/1556-4029.14608. PMID: 33105027.

38. George Mason University. (n.d.). *AI-based bruise detection & analysis for criminal justice applications*. https://ott.gmu.edu/wp-content/uploads/2023/06/JC-GMU-23-018-Lattanzi-Tech-Summary.pdf

39. National Criminal Justice Reference Service. (2023). *Automation-supported curation of large forensic image databases* https://www.ojp.gov/pdffiles1/nij/grants/306559.pdf

40. Cao, J., An, G., Li, J., Wang, L., Ren, K., Du, Q., Yun, K., Wang, Y., & Sun, J. (2023, Mar). Combined metabolomics and tandem machine-learning models for wound age estimation: a novel analytical strategy. *Forensic Sci Res*, 8(1), 50-61. https://doi.org/10.1093/fsr/owad007. PMCID: PMC10265958. PMID: 37415796.

41. Wang, Z., Zhang, F., Wang, L., Yuan, H., Guan, D., & Zhao, R. (2022). Advances in artificial intelligence-based microbiome for PMI estimation. *Front Microbiol*, 13, 1034051. https://doi.org/10.3389/fmicb.2022.1034051. PMCID: PMC9577360. PMID: 36267183.

42. Paba, R., Moniz, R., & Magni, P. A. (2023). Optimizing underwater visual records for crime scene investigations in water with clear to reduced visibility. *Forensic Sci Int Synerg*, 6, 100329. https://doi.org/10.1016/j.fsisyn.2023.100329. PMCID: PMC10220300. PMID: 37249969.

43. Abraham, J., Ng, R., Morelato, M., Tahtouh, M., & Roux, C. (2021). Automatically classifying crime scene images using machine learning methodologies. *Forensic Science International: Digital Investigation*, 39. https://doi.org/10.1016/j.fsidi.2021.301273.

44. Huang, T. Y., & Yu, J. C. C. (2021, Jun 29). Development of crime scene intelligence using a hand-held raman spectrometer and transfer learning. *Anal Chem*, 93(25), 8889-8896. https://doi.org/10.1021/acs.analchem.1c01099. PMID: 34134486.

45. Cooman, T., Ott, C., & Arroyo, L. E. (2023). *Evaluation and classification of fentanyl related compounds using EC-SERS and machine learning*. https://www.ojp.gov/ncjrs/virtual-library/abstracts/evaluation-and-classification-fentanyl-related-compounds-using-ec

46. Koshute, P., Hagan, N., & Jameson, N. J. (2022). Machine learning model for detecting fentanyl analogs from mass spectra. *Forensic Chemistry*, 27. https://doi.org/10.1016/j.forc.2021.100379.

47. Wong, S. L., Ng, L. T., Tan, J., & Pan, J. (2023). Screening unknown novel psychoactive substances using GC–MS based machine learning. *Forensic Chemistry,* 34. https://doi.org/10.1016/j.forc.2023.100499.

48. Tran, T. T. V., Surya Wibowo, A., Tayara, H., & Chong, K. T. (2023, May 8). Artificial intelligence in drug toxicity prediction: Recent advances, challenges, and future perspectives. *J Chem Inf Model*, 63(9), 2628-2643. https://doi.org/10.1021/acs.jcim.3c00200. PMID: 37125780.

49. National Institute of Justice. (2021). *Deep learning to enhance investigative lead information for automotive clear coats.* https://nij.ojp.gov/funding/awards/15pnij-21-gg-04182-ress

50. Wei, C., Wang, J., He, X., & Gu, Y. (2021). A fast and non-destructive approach to identify the heavy mineral oil trace evidence based on spectral fusion treatment and chemometrics. *Microchemical Journal*, 163. https://doi.org/10.1016/j.microc.2021.105924.

51. Swofford, H., & Champod, C. (2022). Probabilistic reporting and algorithms in forensic science: Stakeholder perspectives within the American criminal justice system. *Forensic Sci Int Synerg*, 4, 100220. https://doi.org/10.1016/j.fsisyn.2022.100220. PMCID: PMC8850671. PMID: 35198945.

52. Justia. (2021). *State of New Jersey v. Corey Pickett.* https://law.justia.com/cases/new-jersey/appellate-division-published/2021/a4207-19.html

53. Nutter, P. W. (n.d.). Machine learning evidence: Admissibility and weight. *Journal of Constitutional Law*, 21(3).

54. Johnson, T. L., & Johnson, N. N. (2023). *Police facial recognition technology can't tell black people apart.* https://www.scientificamerican.com/article/police-facial-recognition-technology-cant-tell-black-people-apart/

Forensic Technology
CENTER OF EXCELLENCE

A program of the National Institute of Justice

NIJ.GOV | National Institute of Justice
STRENGTHEN SCIENCE. ADVANCE JUSTICE.

RTI INTERNATIONAL

Published: **December 2023**

Disclaimer

The NIJ FTCOE, led by RTI International, is supported through a Cooperative Agreement from the NIJ (15PNIJ-21-GK-02192-MUMU), Office of Justice Programs, U.S. Department of Justice. Neither the U.S. Department of Justice nor any of its components are responsible for, or necessarily endorse, this in-brief. NIJ is the research, development, and evaluation agency of the U.S. Department of Justice. NIJ is dedicated to improving knowledge and understanding of crime and justice issues through science. NIJ provides objective and independent knowledge and tools to inform the decision-making of the criminal and juvenile justice communities to reduce crime and advance justice, particularly at the state and local levels. The NIJ Office of Investigative and Forensic Sciences (OIFS) is the federal government's lead agency for forensic science research and development. OIFS's mission is to improve the quality and practice of forensic science through innovative solutions that support research and development, testing and evaluation, technology, information exchange, and the development of training resources for the criminal justice community.

Public Domain Notice

Suggested Citation

Shute, R., Gamaleldin, S., Cain, D., Mecray, M., Bollenbacher, J., & Ropero-Miller, J.D. (2023, December). What FSSP leaders should know about artificial intelligence and its role in forensic science. Research Triangle Park, NC: RTI International.

FTCOE Contact

Jeri Ropero-Miller, PhD, F-ABFT
Principal Scientist, FTCOE
jerimiller@rti.org

NIJ Contact

Danielle McLeod-Henning, MFS
Physical Scientist
Office of Investigative and Forensic Sciences
danielle.mcleod-henning@usdoj.gov

Technical Contacts

Rebecca Shute, MS
RTI International
rshute@rti.org

Sara Gamaleldin
RTI International
sgamaleldin@rti.org

Photo Credit:
Page 1—https://www.pexels.com/photo/green-computer-circuit-board-159220/