DECEMBER 2023

Federal PROBATION

a journal of correctional philosophy and practice

Special Issue on Technology and Criminal Justice: Stepping into the Next Stage of Development

Federal PROBATION

a journal of correctional philosophy and practice

PUBLISHED BY
The Administrative Office of the U.S. Courts

Judge Roslynn R. Mauskopf, *Director*

John J. Fitzgerald, *Chief*
Probation and Pretrial Services Office

Federal Probation ISSN 0014-9128 is dedicated to informing its readers about current thought, research, and practice in criminal justice, community supervision, and corrections. The journal welcomes the contributions of persons who work with or study defendants and offenders and invites authors to submit articles describing experience or significant findings regarding the prevention and control of crime and delinquency. A style sheet is available from the editor.

Federal Probation is published three times yearly—in June, September, and December. Permission to quote is granted on the condition that appropriate credit is given the author and *Federal Probation*. For information about reprinting articles, please contact the editor.

Subscriptions to *Federal Probation* are available from the Superintendent of Documents of the Government Printing Office at an annual rate of $16.50 ($22.40 foreign). Please see the subscription order form on the last page of this issue for more information.

Federal Probation can also be accessed online at no charge at www.uscourts.gov.

EDITORIAL STAFF

Charles Robinson, *Executive Editor*
Ellen Wilson Fielding, *Editor*

Federal Probation
Administrative Office of the U.S. Courts
Washington, DC 20544
telephone: 202-502-1651
fax: 202-502-1677
email: Ellen_Fielding@ao.uscourts.gov

Postmaster: *Please send address changes to the editor at the address above.*

THIS ISSUE IN BRIEF

Special Issue on Technology and Criminal Justice: Stepping into the Next Stage of Development

DEPARTMENTS

Introduction to Special Issue on Technology and Criminal Justice: Stepping into the Next Stage of Development

WELCOME TO THIS SPECIAL EDITION of *Federal Probation*. Reflecting on my beginnings as a federal probation officer in the 1980s, I recall the technological landscape being vastly different. Back then, we marveled at beepers, dummy computer terminals, and location monitoring equipment that rivaled the traditional ball and chain in size and utility.

It was hard for me and my colleagues to imagine that, within our careers, we would witness the advent of app-filled cell phones, the Internet, GPS location monitoring tools, and advanced drug and DNA testing. Yet, even more astonishing is the next great leap forward in technology already upon us: the age of Artificial Intelligence and related technologies. Understanding the potential and costs of these technologies is crucial for criminal justice professionals, and that's precisely what this special edition is about.

In "AI in Corrections: The Basics and A Way to Experiment," Dr. Amit Shah, Ashit Chandra, and I offer a foundational understanding of Artificial Intelligence for criminal justice officials who may be navigating this terrain for the first time. We present a practical and impactful use case for the technology related to staff training, which is both secure and avoids ethical pitfalls while being scalable for agencies venturing into the AI landscape.

Thomas H. Cohen's exploration of "The Pretrial Dashboards: Using Technology to Provide Judges with an Understanding of Their Pretrial Release and Detention Decisions" sheds light on the fusion of technology and judicial decision-making. By providing judges with invaluable insights, this paradigm shift not only empowers them with data-driven perspectives but also hints at broader implications for the federal pretrial system and beyond.

In "Exploring Probation and Parole Records Using Natural Language Processing," Hadeel Elyazori, Teneshia Thurman, Kevin Lybarger, and Faye S. Taxman unravel the potential of Natural Language Processing (NLP) to unlock the wealth of information within probation records. Their groundbreaking case study showcases how NLP can revolutionize data interpretation, client management, and policy formulation, heralding a new era of evidence-based practices.

Further delving into the realm of AI, Karine Megerdoomian, Charles E. Horowitz, and Amy B. Marsh underscore the imperative of harnessing advanced technology to address the complex nexus of substance use disorders and mental health conditions in their article on "Automated Extraction of Substance Use and Co-occurring Disorders from Probation Records." They discuss tools now available to automate knowledge discovery from narrative texts, enhancing efficiency and equipping probation offices with actionable insights to navigate multifaceted challenges of rehabilitation and community reintegration.

We close this special edition with "Development and Testing of a Digital Coach Extender Platform for MOUD Uptake" by Jessica Vechinski, Dharmaraj Veeramani, Barbara Bowers, and Todd Molfenter. The authors discuss a pioneering pilot aimed at bridging the gap between the criminal legal system and health systems in combating opioid use disorder. Through the development of a Coaching Extender Platform, this initiative promises to democratize coaching techniques, making them scalable, affordable, and ultimately more impactful in facilitating evidence-based treatments.

I hope you enjoy the articles and find that they spark your interest in the power and proper use of emerging technologies to improve the world of criminal justice and community corrections. From the standpoint of a generation hence, even these cutting-edge tools may appear to be baby steps, as primitive as the beeper in my 1980s probation office.

—Matthew G. Rowland

AI in Corrections: The Basics and a Way to Experiment[1]

Matthew G. Rowland[2]
Amit Shah[3]
Ashit Chandra[4]

THERE IS AN ongoing debate concerning the appropriate role of Artificial Intelligence (AI). This is particularly so in politically charged domains such as criminal justice, where the balance between societal and individual interests is already inherently contentious. As the use of AI moves beyond routine activities and closer to substantive decision-making, such as influencing sentences imposed for criminal violations and how those sentences are enforced, the debate becomes more intense. Often missing in the debate, however, is the perspective of criminal justice officials, including those in community corrections. This omission could be costly, because such officials are uniquely situated to recognize where the technology can enhance operational effectiveness and where it can pose risks to desired outcomes.

Criminal justice officials need to become more familiar with the technology to contribute meaningfully to the debate and its resolution. This can be best achieved through pilots and experimentation specifically designed to surface the technology's strengths, weaknesses, and costs. Without such firsthand experience, criminal justice officials can get lost in the technojargon, hyperbole, and IT industry's self-interest that often hinder a true understanding of AI.

This article aims to provide a basic understanding of AI and the surrounding issues to officials in the criminal justice field, who may not already have this understanding. Also, we advocate for a specific use case for criminal justice agencies to begin their own AI journey. That use case was selected because of its ability to expose the potential and challenges of AI in a relatively safe—yet important—environment. In addition, a successful prototype exists that leverages scalable and economical AI tools, with a criminal justice-oriented thought process already applied. The prototype provides a solid foundation on which to build while reducing time and financial burdens on agencies.

The proposed use case focuses on staff training, which has been identified as a top-tier need of criminal justice agencies (Russo, 2019). At the same time, a training application offers a more controllable environment and involves less sensitive data than would a direct case management use, and allows for extensive human supervision of the AI outputs.

---

[1] This article is the product of human thought and articulation. Generative AI, specifically OpenAI (2021), was used for the quotes in highlighted text boxes. Also, please note Generative AI was provided with a draft of this article and prompted with the question: "Will readers of the Federal Probation Journal find this article interesting." The AI responded: "[. . .] The article covers a critical and emerging topic with a focus on practical application, making it valuable and engaging for readers interested in the intersection of AI and corrections." We'll leave the accuracy of that AI assessment to you, the human reader.

[2] Matthew G. Rowland has more than 35 years of leadership and operational experience in the criminal justice field. He retired from government service as a senior executive and chief of the Probation and Pretrial Services Office for the Administrative Office of the United States Courts. He has since been a principal in Maloney, Rowland, & Associates LLC, providing diverse consulting services to government agencies, think tanks, and global technology firms focused on improving criminal justice outcomes.

[3] Dr. Amit Shah earned his doctorate in Biomedical Engineering at the University of Illinois at Chicago. He led clinical studies on the application of virtual reality and adaptive feedback algorithms on neurorehabilitation of stroke and traumatic brain injury survivors. After earning his Ph.D., Amit has served as a leader of data science initiatives at several companies and has led and won hackathons in predictive analytics. He is an expert on building generative AI solutions, predictive analytics, and What-If modeling. Amit was the Data Science Manager at Abbott Laboratories before leaving to found his own Data and AI consultancy, GNS-AI. He now serves as the President and Chief Data Officer of GNS-AI and is working with various companies and government agencies to deliver safe, cost-effective, and innovative solutions.

[4] Ashit Chandra is Vice President of Infobahn Solutions, specializing in Data Engineering, Data Analytics, and Enterprise Data Management. He has a successful track record designing and implementing Enterprise Data Lakes, Enterprise Data Warehouses, Business intelligence, and Big Data Analytics solutions that empower organizations to harness their data assets and meet mission goals.

An Introduction to AI

AI does not lend itself to easy definition. As noted in a journal from one of the United States' leading technical institutions, "artificial intelligence is constantly evolving, and the term often gets mangled" (Hao, 2018). In an attempt to offer at least something to help readers comprehend the technology, the same article suggests viewing AI as mimicking human intelligence, an analogy frequently employed by others to explain AI as well (Frankenfield, 2023). Unfortunately, the analogy is flawed.

One reason is that human intelligence itself is notoriously hard to define and conceptualize (Weder, 2020). So, the comparison is of one riddle to another. Also, there are clearly fundamental differences in how and why AI and humans operate "cognitively." AI relies on a digital framework, using binary code (a series of zeros and ones) to operate. Further, the technology is directed exclusively by algorithms (programmed step-by-step instructions) that are encoded by humans. Science fiction accounts to the contrary: AI is devoid of consciousness and self-initiative. It does nothing that cannot ultimately be traced back to its human-designed software and hardware. It is an inanimate tool, and its value and impact, whether positive or negative, hinges on the individuals who develop and wield it.

Conversely, humans function within a sensory framework, drawing upon our observations, auditory input, and tactile sensations, all interwoven with our wealth of experience, emotions, and intuition. Human agency and choice are inherent in our decision-making processes. Our collective choice to collaborate, complemented by shared creativity and innovation, gives rise to technologies like AI. It is so sophisticated that it is likely impossible for one person alone to understand it fully. But a full understanding of AI is not what is needed by criminal justice officials.

There are countless examples where we, as individuals, use technology effectively without an in-depth understanding of how it works. Most people would be at a loss to explain the inner workings of their car, cell phones, or personal computers, yet they skillfully use the devices with awareness of what constitutes legal and proper use.

Similarly, when it comes to AI, a conceptual grasp and recognition of appropriate and inappropriate uses are well within the realm of common understanding. While the development and proper use of this technology may necessitate a collaborative effort involving technical, operational, and administrative experts, individuals can ultimately assess the value and correct operation of AI.

What makes AI particularly exciting and disconcerting at the same time is the rate at which it is improving. The fundamental building blocks of AI, including computational speed, the amount of digital data, and sophisticated algorithms, are increasing exponentially (Henshall, 2023)—as are the number and expertise of developers to harness those growing resources. Combining that progress with increased investments in things like robotics conjures images of a dystopian future where computers dominate society, rendering human involvement unnecessary.

However, it is vital to differentiate between the speculative future of AI and its current state and near-term trajectory. The remote and imaginable should not eclipse the present and tangible. As discussed further below in relation to the proposed use case, AI can more than quickly process large stores of digitized text. Through a functionality called Natural Language Programming, it can understand and use human language. Additionally, image analysis functionality can identify objects, faces, scenes, and anomalies in digital photos and videos. It can display the results of its analysis in the manner and context that individual users find most valuable. And it can do all that when designed and resourced correctly in close to real-time, and with a precision and discipline that humans would find hard to match. Think how efficient and effective probation officers and other criminal justice officials would be with AI assistants to take in, process, and display information when and how officials need it. [5]

That alone warrants agencies undertaking greater testing of, and experimentation with, the technology. Furthermore, understanding AI today will better inform us about its future potential and risks. And if those arguments were not enough, then there is the reality (which the field of corrections is a testament to) that others will experiment and use the technology regardless of ethical considerations because of its lucrative potential. If AI is only available to such people, the dystopian future is all but assured.

Nevertheless, there are significant concerns surrounding AI even in its current state, particularly within the criminal justice domain. These concerns include: (1) AI's reliance on historical data and input from existing criminal justice personnel, potentially perpetuating biases, inequities, and inefficiencies attributed to the current system, and (2) the risk of undue deference to technology by future personnel within the criminal justice system (Burns, 2022).

The fear, somewhat paradoxically, is that AI will be overly influenced by poor and bad actors from the past. Conversely, future actors in the criminal justice system will subjugate their own good judgment and overly rely on AI. Fortunately, good technology implementation practices in relation to AI can mitigate those concerns. These practices include transparency and diligent human examination and oversight of AI inputs, outputs, and uses. In addition, the technology itself offers ways to combat faulty inputs and distorted outputs, such as AI-based tests and techniques that can proactively expose potential biases in data analyzed and conclusions drawn (Feast, 2020).

So, well-designed AI applications can actually reduce bias, offering "a number of advantages [over human judgment alone], including the speed at which they process information. Also, because they do not have feelings, they are more objective and predictable than people in their decision-making. They are a core component of overcoming the pervasive bias and discrimination that exists in the criminal justice system" (Rizer & Watney, 2018).

AI Already in Use

It's highly probable that you've interacted with AI today without even realizing it. Whether you used GPS to navigate around traffic, enjoyed personalized music, or engaged with social media, AI played a pivotal role in those experiences. If you're reading this article online, your device, network, and the search engine likely leveraged AI.

While not mainstream, some criminal justice agencies have started using AI in their operations. A consortium organized by the National Institute of Justice found that AI is currently employed to screen prison visitors and incoming mail for contraband. It's also used for analyzing inmate telephone conversations to identify threats and potential

[5] A practical example is that the national and local policies and procedures of the Federal Probation and Pretrial Services System are collectively thousands of pages long. It is impractical to expect individual probation and pretrial staff members to memorize and adhere to so many policy and procedural provisions without support. There are several ways AI can help the officers navigate to the relevant policy guidance when, where, and how they need it.

criminal activity. Additionally, some agencies use chatbots[6] to remind pretrial defendants of court dates and to provide probationers with relevant information to help them comply with their supervision conditions. AI is also sometimes employed for actuarial prediction of recidivist risk presented by inmates and individuals under community supervision.

Although reports on the effectiveness of AI in those instances is not yet publicly available, the consortium concluded that "AI is here to stay" and that "AI-enabled tools have the potential to improve efficiency, reduce costs, and expand capabilities across many criminal justice use cases[.]" To achieve those benefits, the consortium added, "will require intentional investment, careful consideration, and sustained efforts from criminal justice decision makers" (Criminal Justice Testing and Evaluation Consortium, 2020).

A similar conclusion was reached by officials from the Administrative Office of the United States Courts (AO). The AO is a federal judicial agency responsible for overseeing, supporting, and reporting on the Federal Probation and Pretrial Services System (FPPS).

Both the AO and FPPS face the challenge of processing a vast amount of information related to court-involved individuals and the strategies and activities probation and pretrial services officers use in investigating and supervising those individuals.

For probation and pretrial services officers, the task involves sifting through the mass of information provided by clients themselves, the community, and other agencies to identify what is relevant and actionable for effective case management.

For AO administrators, the challenge is identifying systemic patterns and best practices from literally tons of data (if it were printed out) that officers enter into case management systems regarding their clients and the strategies employed to achieve positive case outcomes.

To determine if AI could help meet the challenge of efficiently managing the vast amounts of data to answer operationally important questions, the AO undertook a *proof of concept*. The effort is described in more detail in an *Irish Probation Journal* article (Rowland, Beatty-Gregoire, & Fitzgerald, 2019) but, in short, involved forming two teams, each with a handful of probation officer specialists and computer engineers.[7] Each team was also equipped with open-source[8] AI tools.

One team was provided with the case notes, known as chronological entries, typed by probation officers on 133,000 post-conviction supervision cases; the other group given scanned copies of 11,243 presentence reports. [9]

Each team was then tasked with using the AI tools to answer specific questions.

The team with the supervision case notes was asked to identify specific references in the notes that would justify concluding that the person supervised had ties to violent extremist groups. The team handling the presentence reports had to determine how many defendants, as described in the reports, were suffering from mental illness and the nature of their condition.

Both teams first conducted quality control checks on the documents they were given, standardized the data format to facilitate AI analysis, and developed algorithms to categorize relevant information hierarchically based on the posed questions. They also created output reports that allowed probation officers in the courts, who were familiar with the cases being analyzed, to verify the accuracy of the results. Importantly, the outputs allowed the officers in the courts to see the exact data upon which the AI relied to classify the case as involving persons with ties to violent extremist groups or suffering from mental illness.

Upon reviewing the output reports, the officers in the courts familiar with the cases provided feedback on the reports' accuracy and offered insights into why the results were correct or incorrect in each case. That input led to modifications to the algorithms, and the process was repeated, accuracy improving with each iteration to the point that the outputs were considered highly reliable based on the data analyzed.[10]

The ultimate finding from the proof of concept was that "at roughly 3% of the price of doing it manually and at a fraction of the time, the AI [. . .] revealed insights into violent extremists under supervision and the mental health condition of persons being sentenced in federal court." Further, it was concluded that "[AI] offers unprecedented opportunities to learn from past cases, to make [corrections] more efficient, and to further several public interests."

Consequently, the AO project evaluators recommended additional experimentation with AI—but with the caveat that the agency and those like it considering AI "invest in the front end to ensure business needs are clear and that the AI is properly 'educated' about the data it will be processing. Again, there is strong support for the 'supervised model' of AI with the technology and subject-matter experts working together, rather than independently."

A Proposed New Use Case

Deciding where to begin experimenting with AI can be daunting for any corrections agency. However, it is important to remember that lessons learned from adopting other technological tools can be applied to AI. This includes following generally accepted change-management principles, conducting legal and ethical reviews, developing cost-benefit measures, and not operationalizing anything that could affect actual case management without sufficient testing and vetting. Moreover, technology seldom operates perfectly out of the box and requires ongoing configuration and adjustments for optimization.

As mentioned previously, it is an established best practice to use cross-cutting teams when developing AI applications and to ensure ongoing human supervision of the AI. Such a cross-cutting team can also assist on the front end in defining project goals and thinking about how outcomes can assist in shaping an ultimate vision for AI, assuming the technology proves useful. To that end, to develop a new use case for this article, the first step taken was discussion with federal, state, and local corrections officials and their technology teams.

A vision for the technology that emerged

[6] Chatbots are AI applications that enable technology to engage with humans through speech or text, answering questions, directing queries, and furnishing necessary information.

[7] The computer engineers were a mix of judiciary employees and contract vendors, most with only recent exposure to the AI tools to be used in the project.

[8] Open source refers to software that is publicly available and free to use.

[9] To protect against inappropriate secondary use or disclosure of the case data collected as part of the proof of concept, only government-approved environments were used to store and analyze the data. In addition, judiciary data retention and disposal rules were applied, and all staff involved in the project were subject to confidentiality agreements and government security regulations.

[10] The subject-matter experts and reviewing officers noted that output reports were only accurate to the degree the data analyzed was up-to-date and complete. They noted such analysis is not a "launch and forget" endeavor but rather that "ongoing review of data dictionaries, expansion of data sources and a strong feedback loop with users are needed for the technology to achieve its full potential."

from those discussions was AI as a "digital assistant": not replacing humans but helping improve human decision-making. With AI memory capacity and speedy recall, AI could prove to be a valuable repository of policies, procedures, and best practices and a conveyer of institutional knowledge. With a well-designed interface, AI can give criminal justice officials the information they need, when, where, and how they need it.

As illustration, envision a case that transfers between probation officers. It may not be easy for the new officer to detect changes in the client's appearance or living conditions compared to what occurred before the transfer. In contrast, AI technology can easily compare digitized photos of the client and residence, taken before and after the transfer. The AI could also point out the differences in a relevant way based on what has been learned from other cases, such as distinguishing where weight loss may be a sign of improving health in the client as opposed to resumed drug use or a mental or physical medical problem.

In terms of detecting changes in the client's living conditions, one of the most notorious cases in community corrections history involved a person under supervision who transferred repeatedly among officers and agencies. The transfers contributed to new officers not detecting signs that the person under supervision had modified his house and property to conceal the presence of persons he had kidnapped and whom he repeatedly assaulted during the period of supervision. It is admittedly speculative but interesting to think that AI might have helped the officers detect the changes in the residence over time and led to quicker detection, or ideally deterred the client's criminal behavior.

The vision aside, those in corrections consulted for the identification of a new use case had questions regarding the ethical use of AI. The United States is only beginning to consider regulations pertaining to the technology, with a first-of-its-kind executive order being recently signed by President Biden (White House Briefing Room, 2023). The full impact of that order is not yet known, but it appears in substantive areas to be consistent with regulations a little further along promulgated by the European Union (EU). The EU approach establishes categories of risk based on how and by whom the AI is used, and sets out limitations and requirements commensurate with that risk level.

Under the EU system, correctional agencies' operational use of AI would likely be deemed "high risk." The regulations call for careful thought and documentation related to the data selection, algorithm development, and other inputs into the AI, as well as the technical workings of the technology design itself. The regulations also call for ongoing and rigorous testing and human supervision of the AI outputs and proactive steps to avoid any impermissible biases from influencing the AI and its outputs. An overall requirement under the EU regulations is transparency (European Union, n.d.). Consequently, corrections agencies should keep direct stakeholders informed and consider publishing papers in professional and academic journals about their AI use as well. This will have the added benefit of allowing corrections agencies to learn from each other regarding AI utility and best practices.

With those considerations in mind, the specific use case that we suggest corrections agencies explore to gain familiarity with AI relates to staff training—specifically, using AI to interpret audio recordings of mock interviews between staff and "clients." The training context offers a controlled environment, enabling limitations on sensitive information and identities and plenty of human supervision of the AI outputs. Another factor for the recommendation is that we have successfully developed a prototype that analyzed audio-recorded conversations between probation officers and anonymized or mock clients. For those of you interested in the technical specifications of the prototype, see the end notes.[i]

Beyond its ability to transcribe conversations and identify speakers, the prototype offers both descriptive and qualitative insights into the dialogue. Notably, the ChatGPT-style interface is a compelling feature for extracting this kind of information.

By simply inputting a question like *Who spoke more, Speaker One or Speaker Two?*, the AI promptly responds with the answer. It also allows for more in-depth inquiries, such as "Did any of the speakers use profanity?" or "Did one speaker talk over the other?" or "Did the speakers discuss the facts and circumstances surrounding the client's recent drug use?"

The possibilities are virtually limitless, with the only constraint being the need to develop training materials for the AI. It's worth noting that crafting these training materials demands careful consideration and testing, as their quality significantly impacts the AI's output. For instance, if you intend to assess whether officers in a recorded conversation are using specific techniques, like cognitive-behavioral or motivational interviewing methods, you must carefully define the words, phrases, and even the tone of voice associated with these techniques. Similarly, when gauging the "client's" response to the officers' use of these techniques, you'll need to specify the words, phrases, and tone that the AI should identify.

On the back end, effort was required to provide the AI with feedback regarding the accuracy of its determinations and the rationale behind its decisions. For instance, it was essential to ascertain whether the AI accurately classified the officer's conversational approach as "directive" or "instructional," when in fact, it was more "collaborative" and "emphatic." This feedback loop involved querying the AI through the interface to understand the data it relied on to reach its conclusions. Human supervisors of the application then had to evaluate whether new training material for the technology was necessary or if modifications to the application's algorithm were warranted.

The prototype illustrated the significant potential of audio recording analysis. Words, phrases, their arrangement, and the nuances of tone hold the key to correctly understanding and categorizing a conversation. However, if a jurisdiction so desires, the option of adding visual analysis is available. Given the substantial portion of communication that is non-verbal, supplementing verbal cues with facial expressions, body language, and other non-verbal signals can render the analysis even more comprehensive.

Capturing people's voices and images, however, creates risk. Although in our modern digital world many of our voices and images are floating somewhere in the public domain, what makes certain AI uses, like Deep Fakes,[11] disconcerting is that they can manipulate our voices and images to make it appear we have said and done things we have not. There are some defenses to that, but restricting the data made publicly available helps as well. Consequently, there are privacy considerations that should go into the development of agencies' AI environments. For a brief discussion of such considerations, please see the end notes following the Bibliography.[ii]

Bibliography

Burns, G. (2022, February 1). *The use and future*

[11] Deep Fakes refer to images created or manipulated using AI technologies for purposes of deceiving those that view it. See, https://www.forbes.com/sites/alexandralevine/2023/10/12/in-a-new-era-of-deepfakes-ai-makes-real-news-anchors-report-fake-stories/?sh=36960f2957af

of artificial intelligence monitoring in prisons. Retrieved from The Reasons Foundation: https://reason.org/commentary/the-use-and-future-of-artificial-intelligence-monitoring-in-prisons/

Criminal Justice Testing and Evaluation Consortium (2020, August). *Artificial intelligence applications in corrections.* Retrieved from National Institute of Justice, U.S. Department of Justice: https://cjtec.org/files/64bfb2359c420

European Union. (n.d.). *Shaping Europe's digital future*. Retrieved March 17, 2024, from European Commision: https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai#:~:text=The%20AI%20Act%20allows%20the,EU%20fall%20into%20this%20category.

Feast, J. (2020, October). *Root out bias at every stage of your AI-development process.* Retrieved from Harvard Business Review: https://hbr.org/2020/10/root-out-bias-at-every-stage-of-your-ai-development-process

Frankenfield, J. (2023, April 24). *Artificial Intelligence: What it is and how it is used.* Retrieved from Investopedia: https://www.investopedia.com/terms/a/artificial-intelligence-ai.asp

Hao, K. (2018, November 10). *What is AI? We drew you a flowchart to work it out.* Retrieved from MIT Technology Review: https://www.technologyreview.com/2018/11/10/139137/is-this-ai-we-drew-you-a-flowchart-to-work-it-out/

Henshall, W. (2023, August 2). *4 charts that show why AI progress is unlikely to slow down.* Retrieved from Time: https://time.com/6300942/ai-progress-charts/

Rizer, A., & Watney, C. (2018). Artificial intelligence can make our jail system more efficient, equitable, and just. *Texas Review of Law and Politics*, 181-227, 183.

Rowland, M. G., Beatty-Gregoire, N., & Fitzgerald, J. J. (2019, October). Testing artificial intelligence in the United States Probation and Pretrial Services System. *Irish Probation Journal*, 107-117, 116. Retrieved from https://www.probation.ie/EN/PB/0/D9C80060AC8ED236802584C100510922/$File/Testomg%20Artificial%20Intelligence%20in%20the%20United%20States%20Probation%20and%20Pre-Trial%20Services%20System.pdf

Russo, J. (2019, December 1). *Workforce issues in corrections.* Retrieved from National Institute of Justice: https://nij.ojp.gov/topics/articles/workforce-issues-corrections

Tim Fountaine, B. M. (2019, July). *Building the AI-Powered organization.* Retrieved from Harvard Business Review: https://hbr.org/2019/07/building-the-ai-powered-organization

Vynck, G. D. (2023, August 29). *AI images are getting harder to spot. Google thinks it has a solution.* Retrieved from *Washington Post*: https://www.washingtonpost.com/technology/2023/08/29/google-wants-watermark-ai-generated-images-stop-deepfakes/

Weder, A. (2020, October 5). *Q&A – What Is Intelligence?* Retrieved from News & Publications: John Hopkins School of Medicine: https://www.hopkinsmedicine.org/news/articles/2020/10/qa--what-is-intelligence

White House Briefing Room. (2023, October 30). *White House.* Retrieved from Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence: https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/

End Notes

i. For the proof of concept centered around interview analysis via generative AI, we employed a carefully curated tech stack to maximize efficiency and performance. Streamlit was our choice for web app development due to its interactive and straightforward nature, making it ideal for rapid prototyping. We harnessed the power of OpenAI's GPT-3.5 Turbo model through their API for sophisticated language modeling, analysis, and summarization, ensuring in-depth insights from the interview data. We also leveraged OpenAI's Whisper model to create a transcript that GPT-3.5 turbo could process. The ChromaDB, an open-source vector database, facilitated efficient data management and indexing to allow for inferencing on the transcript, while LangChain was invaluable for creating a seamless chain of prompts, enhancing the user interaction and data input process, which resulted in a chatbot we could use to dynamically query the resulting transcript. All of this was adeptly put together using Visual Studio Code as the integrated development environment (IDE) for its versatility and extensive developer support.

ii.To protect sensitive audio data and transcripts, for example, a multi-layered on-premise security approach should be taken. The audio files should be transcribed using private voice-to-text models that are served locally and private large language models that are trained internally using the organization's data. This prevents exposing the raw audio to external cloud services. The audio and resulting transcripts should be encrypted and anonymized to remove identifiers. Any additional natural language processing or machine learning inferencing on the transcripts should occur locally on private edge servers, not in the cloud. Strong access controls and audit trails should track all data access, with logs monitored for unauthorized usage. The data should be stored on local servers with hardened security including firewalls, intrusion prevention, and minimal ports exposed. Regular pen testing should check for vulnerabilities. With proper encryption, private models, access controls, on-premise infrastructure, and auditing, the confidentiality of the audio data can be maintained from transcription through usage.

The Pretrial Dashboards: Using Technology to Provide Judges with an Understanding of Their Pretrial Release and Detention Decisions

Thomas H. Cohen
Senior Social Science Analyst
Probation and Pretrial Services Office

WHEN A PERSON (i.e., a defendant) is charged with committing a federal offense, judicial officials have the discretion to determine whether that defendant should be released pretrial, subject to the criteria required by the Eighth Amendment and under 18 U.S.C. §3142 of the federal statute. Under both guiding documents, detention is reserved only for rare cases where "no condition or combination of conditions will reasonably assure the appearance of the person as required and the safety of any other person and the community" (see 18 U.S.C. §3142). The decision to release a defendant into the community or detain the defendant until the case is disposed is of crucial importance. Not only can a defendant's liberty, and therefore, constitutional rights, be constrained by the detention decision, but research has shown that subsequent case outcomes (including the likelihood of conviction, severity of sentence, and long-term recidivism) can be negatively affected when pretrial detention is mandated (Gupta et al., 2016; Heaton et al., 2017; Oleson et al., 2014).

Despite the crucial, some would even say pivotal, role (Carr, 2017) of the pretrial release decision in the federal system and the various provisions and efforts aimed at reducing unnecessary detention, the federal pretrial detention rate remains at a level that has been viewed as high and a source of concern. For example, the percentage of defendants released pretrial (excluding undocumented non-citizens) has declined from 55 percent in fiscal year 2008 to 47 percent in fiscal year 2017 (Cohen & Austin, 2018). Since 2017, the release rate for defendants who are not undocumented non-citizens has remained relatively stable; in fiscal year 2022, for example, the release rate for these defendants was 47 percent (AO, Table H-14B).

In response to these concerns about increasing rates of pretrial detention, the Probation and Pretrial Services Office (PPSO) of the Administrative Office of the U.S. Courts (AO) was tasked with developing a series of statistical dashboards that would allow judges to view their own pretrial release rates by a variety of characteristics and compare them to the nation or their circuit or the district where they preside. These dashboards were created and then disseminated to the federal judicial community in early 2022 and updated since then. Since their release, judges have had ready access to their release and detention decisions for the first time. Before the advent of the pretrial dashboards, this information for the most part was not readily available to judicial officials; rather, judicial officials who were interested in reviewing their release and detention decisions had to rely upon data manually compiled for them by U.S. probation or pretrial services officers within their districts.

This article will provide an overview of the pretrial dashboards that have been created for federal judges, including 1) background about the processes that led to the creation of the dashboards, 2) specific examples of information made available to judges through the dashboards, 3) trainings that have been conducted to introduce judges to the dashboards and the potential impacts of training on dashboard usage, and 4) future implications of the dashboards for the federal pretrial system and the potential of these dashboards to be further disseminated to the public.

Pretrial Dashboards Background

The pretrial dashboards were initially developed in response to requests by judges from the Magistrate Judges Advisory Group (MJAG) and other judicial entities (e.g., Criminal Law

Committee) for a statistical tool that would allow judges to examine their own decision-making on pretrial release. Moreover, these dashboards were intended to further the requirement in 18 U.S.C. § 3154(9) that pretrial services "develop and implement a system to monitor and evaluate bail activities, provide information to judicial officers on the results of bail decisions, and prepare periodic reports to assist in the improvement of the bail process." It was also anticipated that the dashboards would provide a tool for judges as well as probation/pretrial chiefs to monitor release rates and encourage dialogue aimed at reducing various forms of unnecessary detention.

Before the advent of the dashboards, judges did not have the capacity to readily examine their own pretrial release and detention decisions. There was no systematic way for judges to determine the number and percentage of defendants they released pretrial, the extent to which their release decisions varied by key characteristics (e.g., most serious offense charge, pretrial risk assessment (PTRA) risk scores, demographic characteristics), and the rates at which those they placed on release engaged in such pretrial misconduct as missing their court appearances, having an arrest for new crimes, or being revoked on technical violations. Any judge interested in reviewing this information would have to manually collect pretrial data about defendants appearing before their court, a time consuming and laborious process.

The dashboards address these informational gaps for the first time by providing a myriad of pretrial metrics through an interactive format. Specifically, judges can use these dashboards to explore their own pretrial release decisions, ascertain how these release decisions vary by certain criteria (such as PTRA risk scores, most serious conviction offenses, and demographic characteristics), and determine how many defendants they release commit pretrial violations (pretrial rearrest, failure to appear (FTA), or revocation). Judges can also use this information to compare their decisions with the release patterns manifested at the national level or in the circuit/district where they work.

Construction of the Pretrial Dashboards

The pretrial dashboards were constructed through a two-stage process. Initially, the raw pretrial data were obtained from the AO's Probation and Pretrial Automated Case Tracking System (e.g., PACTS). These data were then exported to the Tableau software platform, which provides users with the capacity to create and display interactive analytics. A series of dashboards were constructed and reviewed by subject matter experts within PPSO, who provided crucial assistance and advice about the dashboards' content and graphical design. The dashboards have since been reviewed by several oversight committees, including the Magistrate Judges Advisory Group, Criminal Law Committee, and senior executive staff with PPSO and the AO, who provided additional suggestions and comments.

The dashboards contain information on pretrial activations encompassing ten-year time frames. The initial series of dashboards disseminated to the Judiciary in 2022 included pretrial activations between fiscal years 2011 through 2020, while the 2023 update included pretrial activations that took place between fiscal years 2012 through 2021. The dashboards will be refreshed again in 2024; when this occurs, the dashboards will contain pretrial activations for fiscal years 2014 through 2023. While the dashboards include relatively recent pretrial data, it is important to acknowledge that they do not provide real-time data on judicial release and detention decisions. Hence, judicial officials and other users may decide to review them intermittently, because they remain unchanged for periods spanning 12 months.

During the construction of these dashboards, several limitations were placed on them that should be noted. First, it is crucial to acknowledge that the dashboards were built to enable judges to view their own release decisions but not those of other judges. In other words, judges are unable to use these dashboards to examine and inspect the decisions of other judges within their district or in other districts. Second, federal probation and pretrial services officers are not provided with access to the dashboards at this time because of concerns that, by highlighting the historical release practices of individual judges, the dashboards might hinder officers from making independent release and detention recommendations. It was, however, agreed that chief and deputy chief probation and pretrial services officers would be provided with judge-identifying release and detention information, because these officials were best positioned to work with judges on ways to reduce unnecessary pretrial detention and are statutorily mandated under 18 U.S.C. §3154(9) to provide information and periodic reports to judicial officers that assist in the improvement of the bail process. The probation/pretrial chiefs and deputies can only examine judge-specific data within their own districts; they are precluded from viewing the decisions of judges in another district. Last, demonstrations of the dashboards were provided to officials within the U.S. Department of Justice and the Federal Defenders Office; both entities expressed interest in having a modified version of the dashboards, without any judge-specific information, made available to them through the U.S. Courts website.

An Example of the Pretrial Dashboards

This section provides visual examples of the dashboards through a series of screenshots. The first screenshot shows what the typical dashboard looks like. Specifically, this dashboard presents information on yearly release rates in two fields. The upper field provides national-level yearly release data, while the bottom field displays yearly release data for a particular judge whose name has been deidentified. A judge examining these dashboards can see how many defendants that judge had released for a period spanning fiscal years 2012 through 2021 and, importantly, compare those release rates to those of the nation. (See Figure 1.)

The next screenshot demonstrates the interactive nature of these dashboards. This example illustrates a judge's ability to select certain criteria using various filters placed on the dashboard's right side. In this instance, the application of these filters allows judges to review their release outcomes for only U.S citizens defendants. Note that the filter applies to both data panels, meaning that the national- and judge-level release rates have been filtered to include only U.S. citizen defendants. Undocumented and documented non-citizens and persons of unknown citizenship have been removed from the dashboards. (See Figure 2, page 12.)

The next screenshot further highlights the types of filters available on the pretrial dashboard tool. In this screenshot, the release rates have been further filtered to include only U.S. citizen defendants with cases activated in the Eleventh Circuit, where this judge hears cases. For this dashboard, the release rates have been further adjusted so that the upper data panel reflects the release rates for defendants with cases activated in the Eleventh Circuit. (See Figure 3, page 13.)

Another example of the interactivity of

these dashboards is shown in the next screen shot. Here cases have been further filtered to reflect pretrial activations involving U.S. citizen defendants charged with drug offenses. (See Figure 4, page 14.)

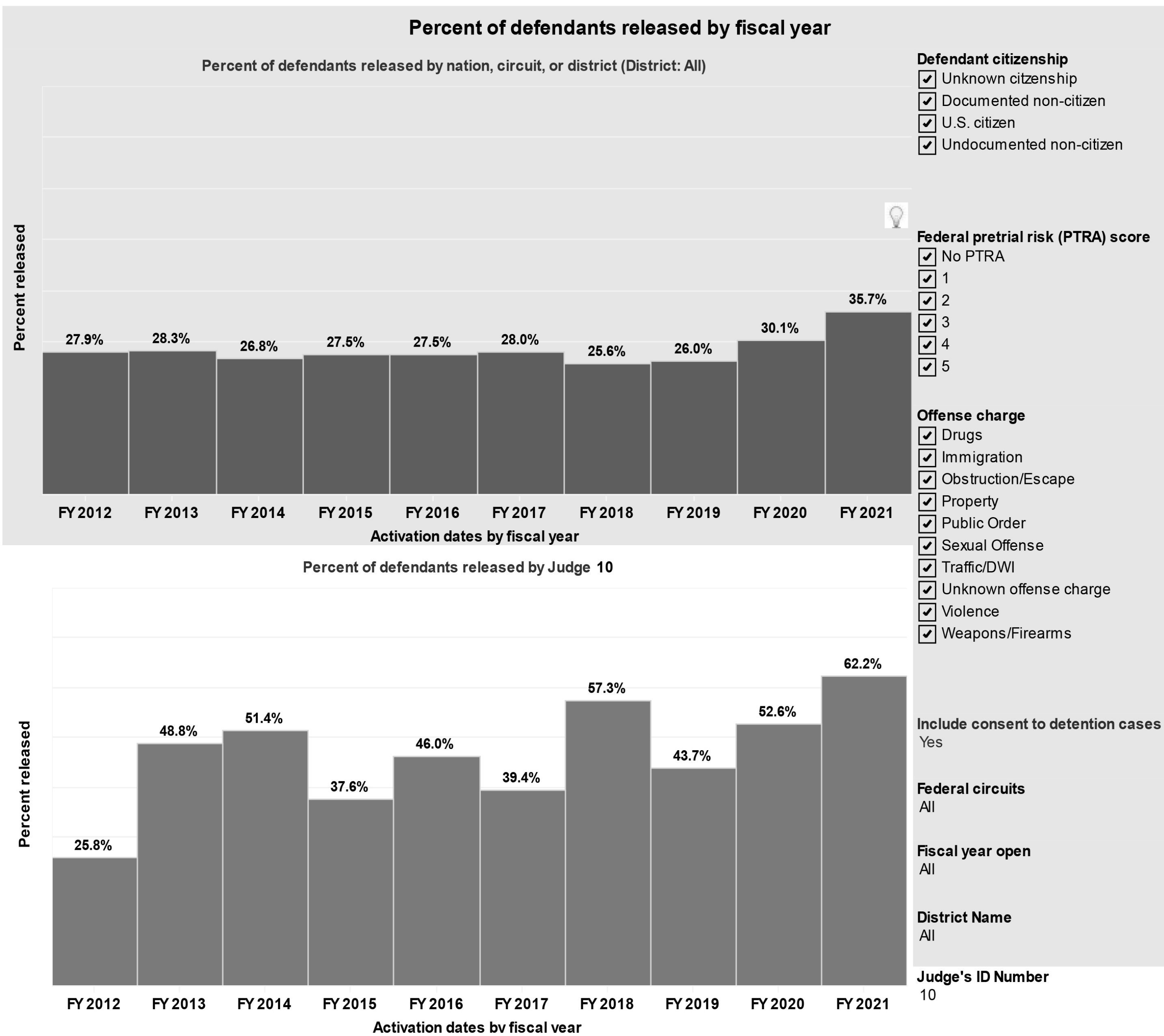
It should be noted that other filters could be applied to these dashboards. For example, users could employ filters encompassing the PTRA risk score, consent to detention cases, and district of case activation to further refine these pretrial release data.

In addition to highlighting yearly release rates, the dashboards contain a variety of other pretrial metrics, some of which are showcased in this article. For example, judges can use the dashboards to examine their release rates by the PTRA's five risk categories.[1] As shown in the screenshot below, the dashboards show release rates declining in a stepwise manner by the five PTRA risk categories both nationally and for this specific judge. (See Figure 5, page 15.)

Another dashboard provides information on release rates by the most serious offense charge both nationally and at the judge level (see next screenshot). As shown, at the national level defendants charged with traffic/DWI, property, or public-order offenses had the highest release rates, while defendants with violence, weapons, or unknown offense charges were the least likely to be placed on pretrial release. Also, all non-citizen defendants (documented or undocumented) have been filtered out of this dashboard. If the non-citizens had been included, then defendants charged with immigration offenses would

[1] The PTRA is an actuarial risk tool used by the federal system to classify defendants by their probability of failure while on pretrial release. The PTRA places defendants into one of five risk categories; the higher the risk grouping, the more likely according to the PTRA that a defendant will fail (that is be rearrested, miss a court appearance, or be revoked), while on pretrial release. See Cohen and Lowenkamp (2019) for an overview of the PTRA risk tool and its capacity to predict pretrial violations.

FIGURE 1



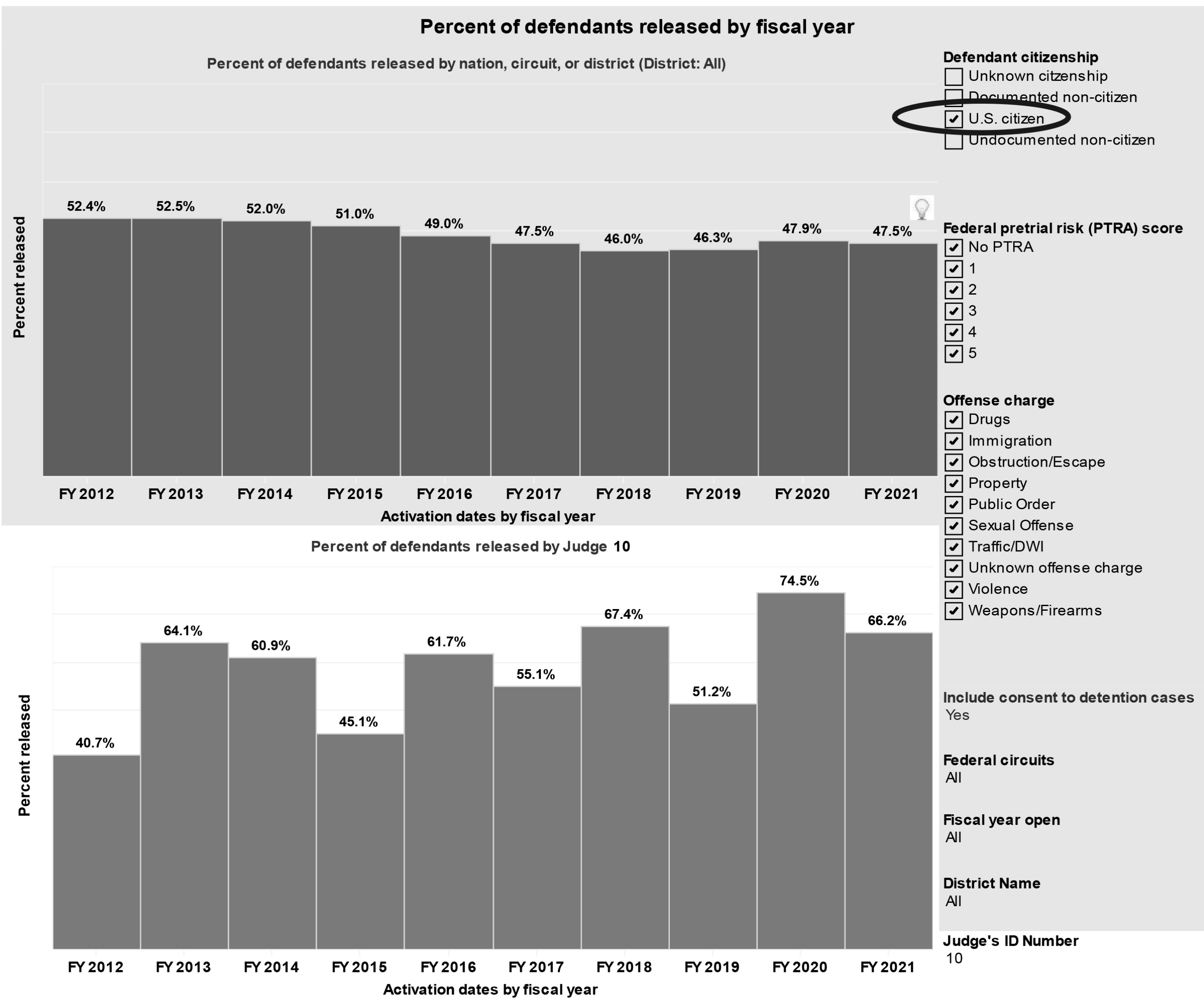
have the lowest rates of pretrial release (data not shown). (See Figure 6, page 16.)

Another dashboard highlighted in this article illustrates this tool's capacity to provide judges with information on how defendants are being detained pretrial. The above data panel provides detention type information, filtering out non-citizen defendants, while the below data panel highlights detention information for a specific judge (again filtering out non-citizen defendants). For the detention dashboard, note that a sizable percentage of detained defendants (43 percent) consented to being detained pretrial. Again, note that users can apply a variety of different filters that would allow them to compare the mechanisms they use for detention with national-, circuit-, or district-level data. (See Figure 7, page 17.)

The last dashboard highlighted in this article showcases how judges can use these tools to better understand the violation rates among their released defendants and examine how these rates vary by the PTRA risk categories. As with the other dashboards, the upper data panel provides national-level information on the percentage of released defendants who were revoked, rearrested, failed to appear (FTA), or had a rearrest for a violent offense across the five PTRA risk categories. Similar to the other dashboards, users could filter out certain case types or assess the violation patterns at the circuit or district level. The below data panel provides information on violations for a specific judge, which is crucial, because judges can now ascertain of those defendants they release how many were rearrested, failed to appear, or had a pretrial revocation by the five PTRA risk categories. (See Figure 8, page 18.)

While this article provides a general overview of the types of data available in these dashboards, it should be stressed that not all data metrics could be highlighted. Specifically, dashboards have also been generated that allow users to compare release rates across the federal judicial districts, highlight release recommendations by pretrial officers and U.S. attorneys, assess release decisions by a defendant's demographic characteristics (e.g., race/

FIGURE 2



ethnicity and gender), and provide details on the average number of special conditions (such as substance abuse testing and location monitoring) imposed on release defendants.

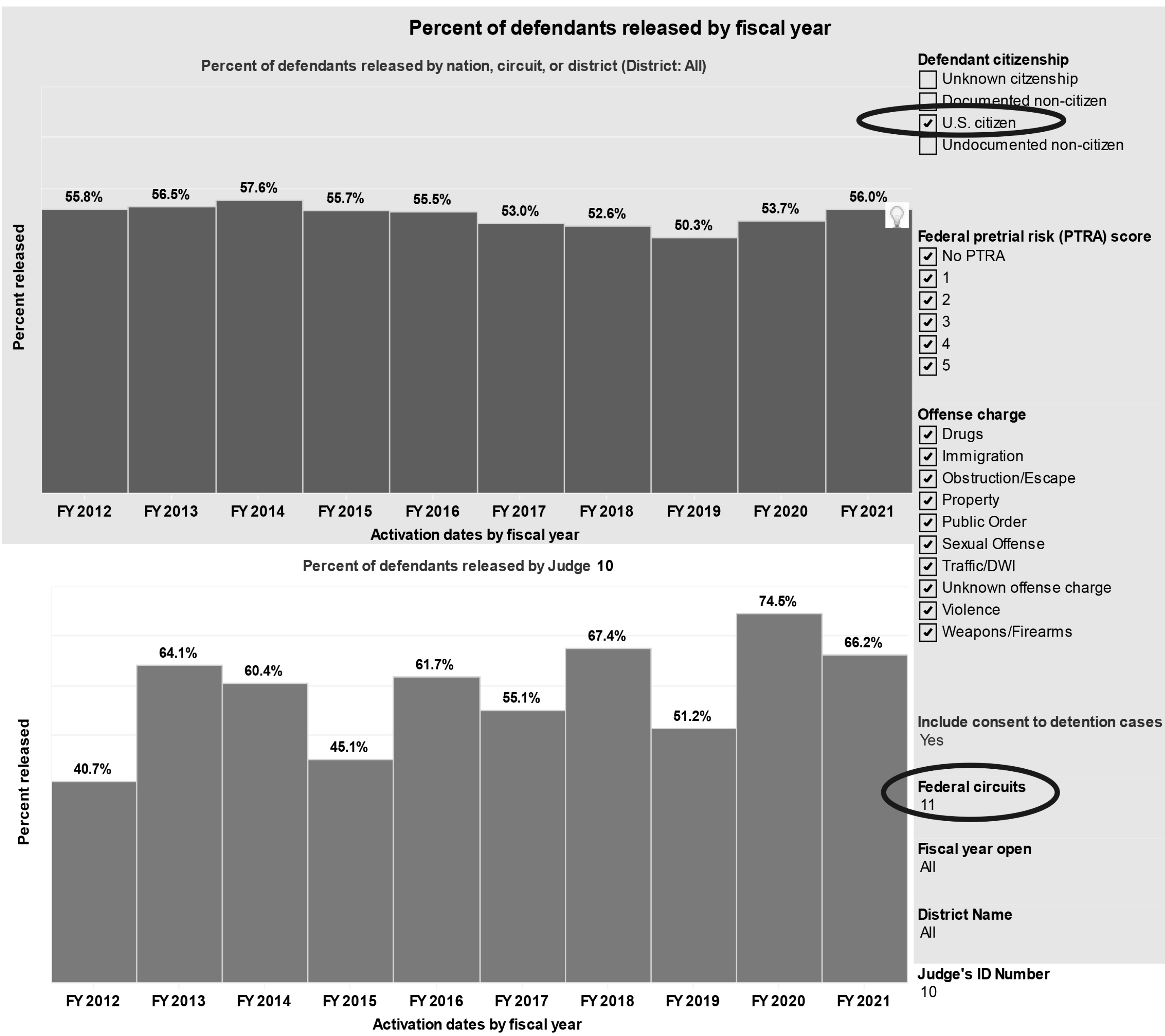
Dashboard Usage and Trainings

While the dashboards provide judges with a plethora of detailed information about their pretrial release and detention decisions, the overall use of these dashboards has been somewhat limited. During the period encompassing the most recent dashboard data update (late February 2023) and the time that this article was written (early August 2023), a total of 100 magistrate judges, representing 16 percent of all full-time federal magistrate judges, viewed the dashboards at least once.[2] Among those judges using the dashboards, 46 percent viewed the dashboards 10 times or more, while 12 percent viewed them only once. Although the dashboards were accessed over 1,000 times on the date that notification of the update occurred—February 17, 2023—since that time, dashboard usage has ranged from 0 to 46 views per day; on most days, the dashboards were accessed an average of about 13 times per day (average was calculated by omitting February dates).

[2] Article III judges were not included in the usage calculations, because for the most part these judges are not involved in the decision to release or detain federal defendants.

The extent to which probation and pretrial chiefs and their deputies and assistant deputies are using the dashboards since February 2023 release date has also been tracked. Of the 300 probation and pretrial chiefs and deputies with access to the dashboards, a total of 52 chiefs, deputies, and assistant deputies from 40 districts viewed the dashboards from 1 to 54 times. These chiefs, deputies, and assistant deputies accounted for about 17 percent of personnel with access to the dashboards. From February 2023 until early August 2023, the daily usage for chiefs and deputies ranged from 1 to 69 views; on average, the dashboards were accessed by chiefs, deputies, and assistant deputies about 11 times per day.

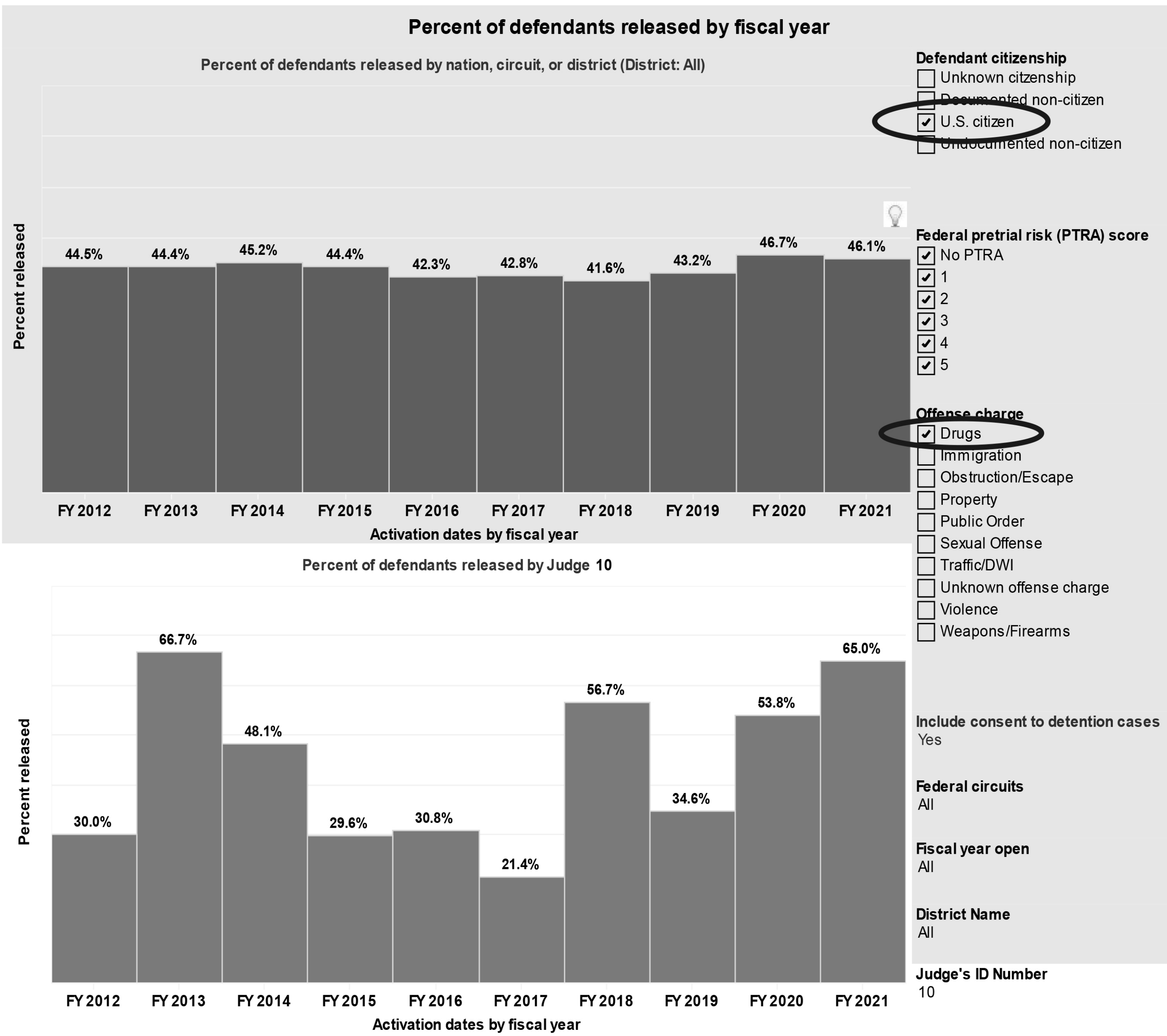
FIGURE 3



To further disseminate information about the pretrial dashboards to the federal judiciary and potentially increase their overall usage, PPSO, in collaboration with the Federal Judicial Center (FJC), engaged in two national trainings aimed at educating judicial officials about these dashboards. The trainings were conducted in April and July 2023 and encompassed background information about the dashboards, instructions on how to access them, and details on the various pretrial metrics available through these dashboards and their capacity to illuminate judicial-level release and detention decisions. After these trainings, an examination of the number of times judges accessed the dashboards was conducted. While there was some increase in dashboard usage around the training periods, the spikes in dashboard access were relatively short and did not differ appreciably from other dates where spikes in dashboard use occurred. (see Figure 9, page 18.)

In addition to these national-level trainings, several localized workshops aimed at introducing judges to the dashboards were conducted. These workshops were part of a larger program being implemented at the district level aimed at reducing unnecessary pretrial detention. In the districts where these localized trainings took place, dashboard usage was examined before and after they occurred. After the trainings, some judges in these districts made more extensive use of the dashboards. For example, in one district two magistrate judges who had not previously used the dashboards began to make extensive use of them after the training; however, the remaining eight judges in this district did not manifest extensive dashboard use. In another example, 3 of the 17 judges made greater use of the dashboards after a training occurred; however, the remaining 14 judges did not use the dashboards more extensively. Last, in a remote training in a district involving relatively few magistrate judges, dashboard use increased for those judges who attended the training workshop.

FIGURE 4



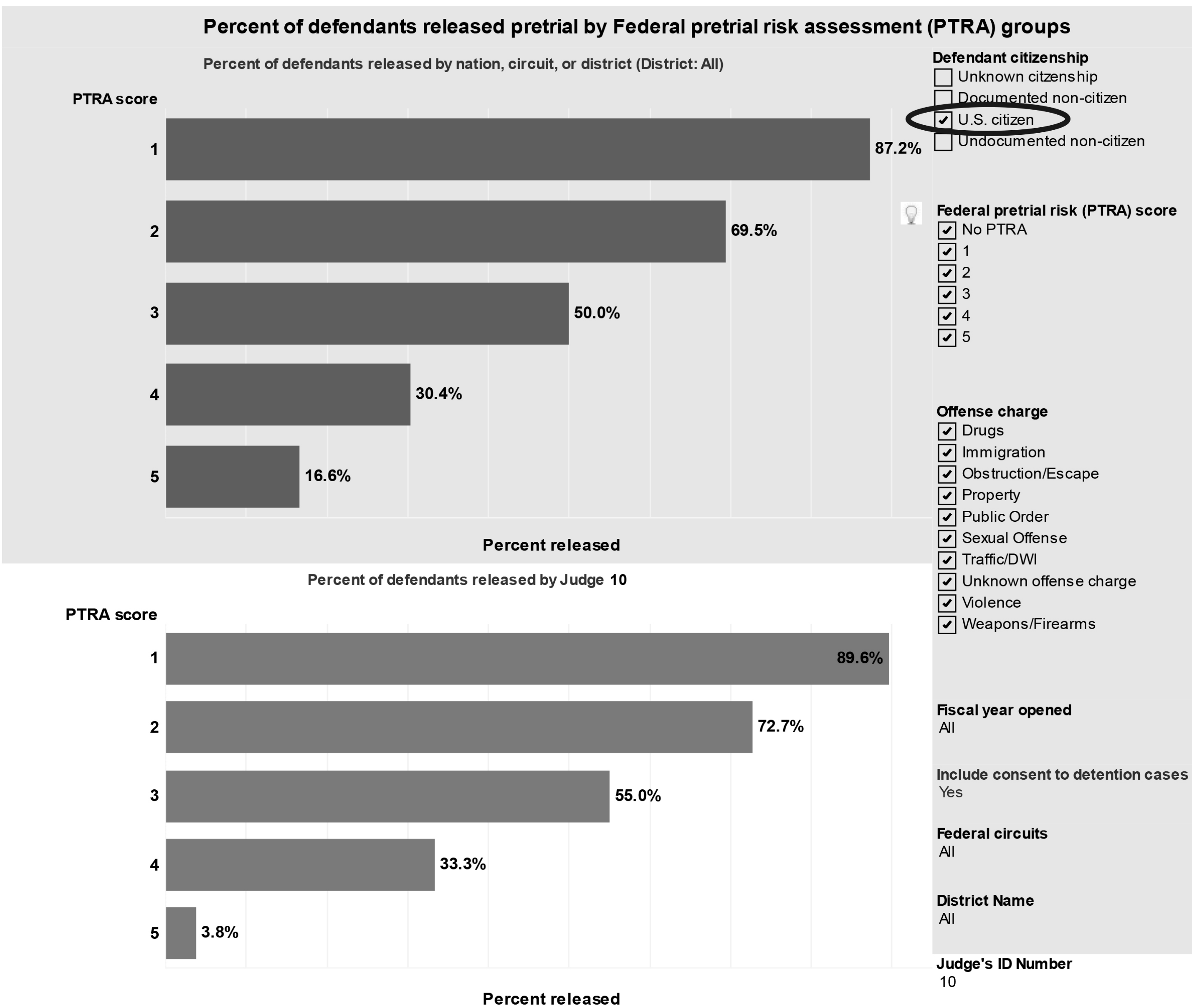
Conclusion and Future Implications for Dashboards

In early 2022, PPSO deployed a series of dashboards that provided judges for the first time with the capacity to examine a wealth of pretrial information to which judicial officers previously had limited access. Specifically, judges can now use these dashboards to examine their own pretrial release patterns and assess the relationships between pretrial decision-making and several factors associated with release (e.g., most serious offense charges, PTRA risk categories, demographic characteristics, citizenship, etc.). Importantly, judges can use these dashboards to compare their release decisions to pretrial outcomes at the national level or the circuit or district where they preside at court. In addition to providing release and detention metrics, these dashboards illuminate information on the types of mechanisms used to detain defendants, the average number of special conditions imposed on released defendants, and the percentage of released defendants who violate their pretrial terms by being rearrested, missing court appearances, or having their release status revoked.

With the advent of the pretrial dashboards, judges now have direct access to data allowing them to analyze their pretrial release and detention decisions. Although the pretrial dashboards provide ready access to data, unfortunately, their use has not been as extensive as initially anticipated. Over the several months previous to the writing of this article, less than a fifth of all magistrate judges and all probation and pretrial chiefs, deputies, and assistant deputies have accessed these dashboards at least once. These results might have occurred because the dashboard tools are still relatively new to the federal judiciary; perhaps more time is required to acclimate judges and pretrial/probation staff to these interactive systems.

To promote further use of the dashboards, PPSO will continue updating the dashboards yearly; however, in the next dashboard refresh, the dashboards will no longer be providing

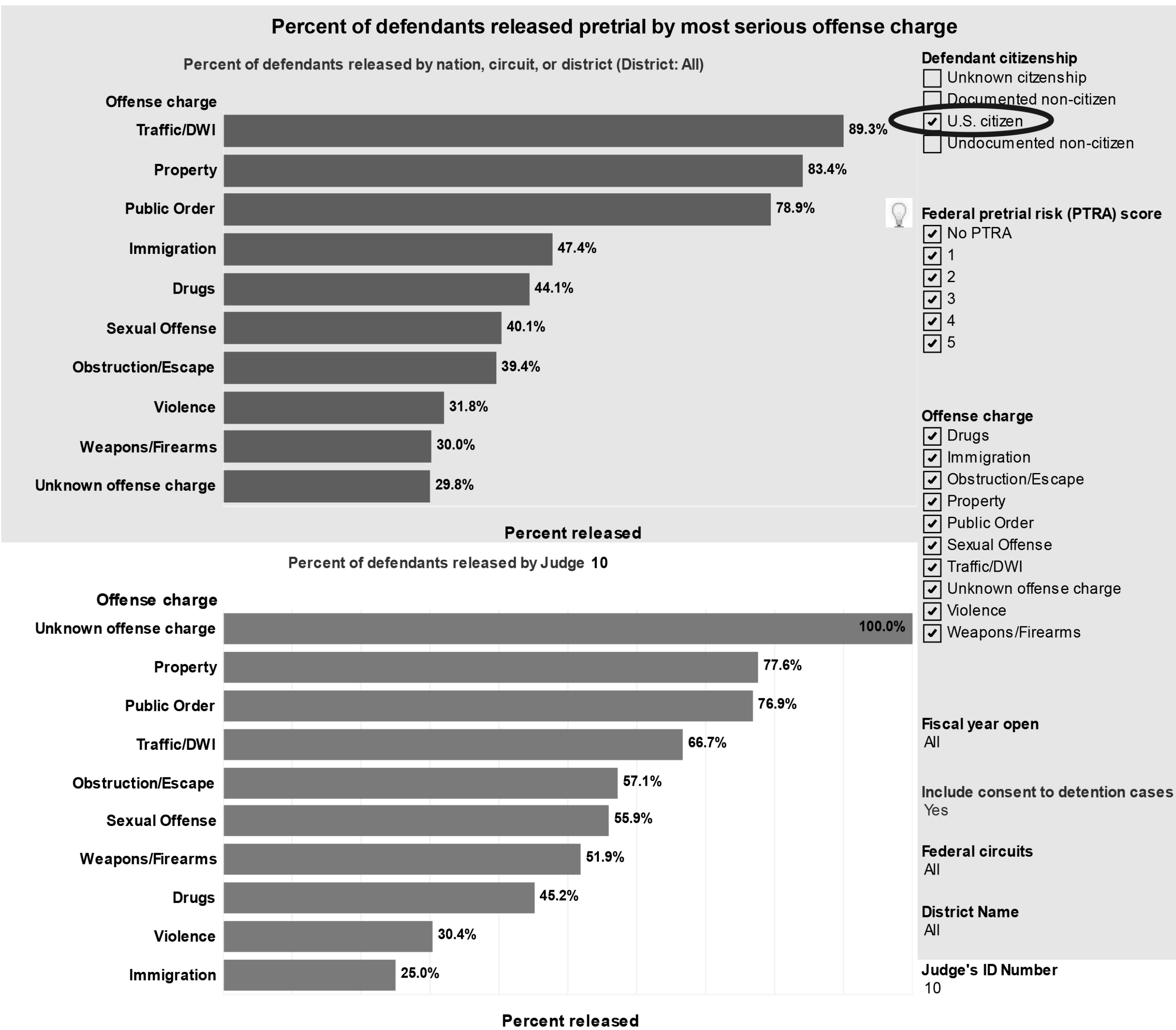
FIGURE 5



data that is one year behind the current fiscal year. Instead, after the update takes place, the dashboards will include pretrial activation data for the ten-year fiscal time frame between 2014 and 2023. Continued judicial officer training (virtual and in person) and outreach is recommended to enhance use. Previous trainings have focused on providing judges with an overview of how the dashboards can be used to illuminate pretrial decision-making in their districts, and subsequent trainings will continue to advocate for their increased use. Training and outreach on the intricacies of these dashboards can also be provided to probation/pretrial chiefs, deputies, and assistant deputies. Beyond training and outreach efforts, PPSO will need to gather feedback from the MJAG, judicial officers, and the probation and pretrial community to solicit their thoughts and suggestions on ways to make the dashboards more relevant to the judicial community.

In addition to these efforts, PPSO has considered the importance of making a modified version of these dashboards available to the public. Nationally, the effort to reduce unnecessary pretrial detention requires collaboration across various stakeholder groups. Prior to the initial release of these dashboards, overviews of these tools were provided to several officials within the U.S. Department of Justice and the Federal Defenders Office. Both entities expressed interest in having the dashboards made available to the public on the uscourts.gov website. This version, unlike the ones currently accessible by judges and probation/pretrial staff, would not contain judge-specific release and detention information; however, it would contain national-, circuit-, and district-level pretrial release data that could be viewed through a variety of interactive filters. In addition to allowing prosecutors and defenders access to these crucial pretrial data, a variety of judicial officials, including Article III judges and newly appointed magistrates, would have access that they currently lack because they hear relatively few federal pretrial cases (e.g., Article IIIs) or (in the case of recently seated magistrates) because they have not been in the system for enough time for their cases to be included in the dashboards. Having a publicly available

FIGURE 6



series of modified pretrial dashboards would provide these judicial officials with the capacity to access these data. Last, researchers, policymakers, and the public could use these dashboards to attain a better understanding of the federal pretrial system.

The pretrial dashboards are a crucial instrument that federal judges can use to understand their release decisions, compare these decisions to national-, circuit-, and district-level data, and assess the extent to which certain types of factors (such as most serious offense charge, PTRA risk score, and race/ethnicity) are associated with release rates. Moreover, these dashboards provide judges with an opportunity to examine other pretrial metrics, including types of detention, special conditions imposed, and instances in which those released are rearrested for new crimes, revoked, or fail to appear. In addition to making these key pretrial data available to judges, PPSO has provided probation and pretrial chiefs and their deputies with dashboard access to encourage further dialogue with judges on ways of ameliorating unnecessary detention. While the dashboards are available only to judges and probation/pretrial chiefs and deputies at this time, we hope that eventually they will be released in a modified form to a larger audience of persons with an interest or stake in the federal pretrial system.

References

AO. (2022, September 30). *Table 14-B: Pretrial services release and detention—excluding illegal alien cases – for the 12-month period ending September 30, 2022.* Retrieved August 2, 2023, from https://www.uscourts.gov/statistics/table/h-14b/judicial-business/2022/09/30.

Carr, J. G. (2017). Why pretrial release really matters. *Federal Sentencing Reporter, 29*(4), 217–220.

Cohen, T. H., & Austin, A. (2018). Examining federal pretrial release trends over the last decade. *Fed. Probation*, *82*, 3.

Cohen, T. H., & Lowenkamp, C. T. (2019). Revalidation of the federal PTRA: Testing the PTRA for predictive biases. *Criminal Justice and Behavior, 46*(2), 234–260.

FIGURE 7

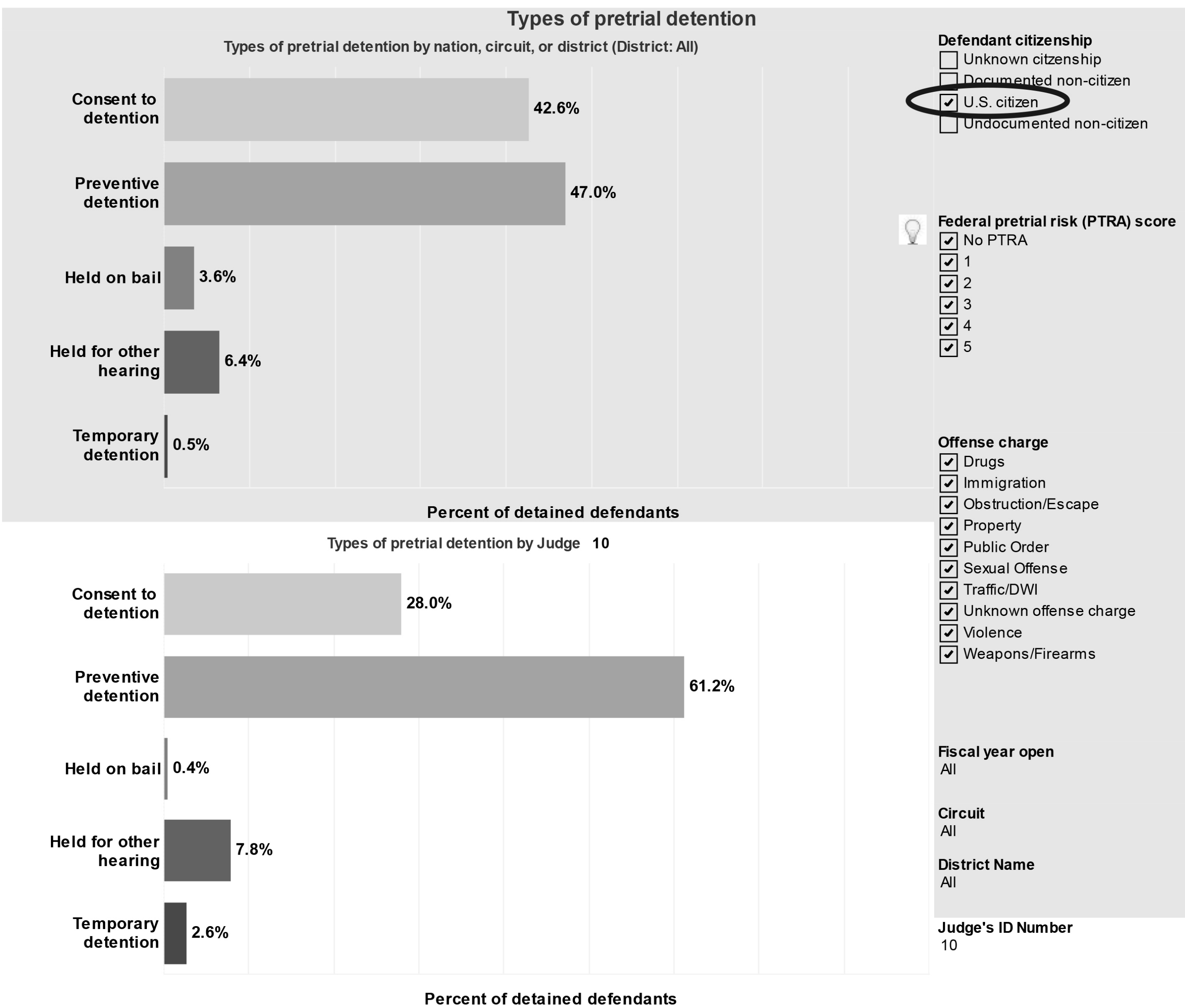
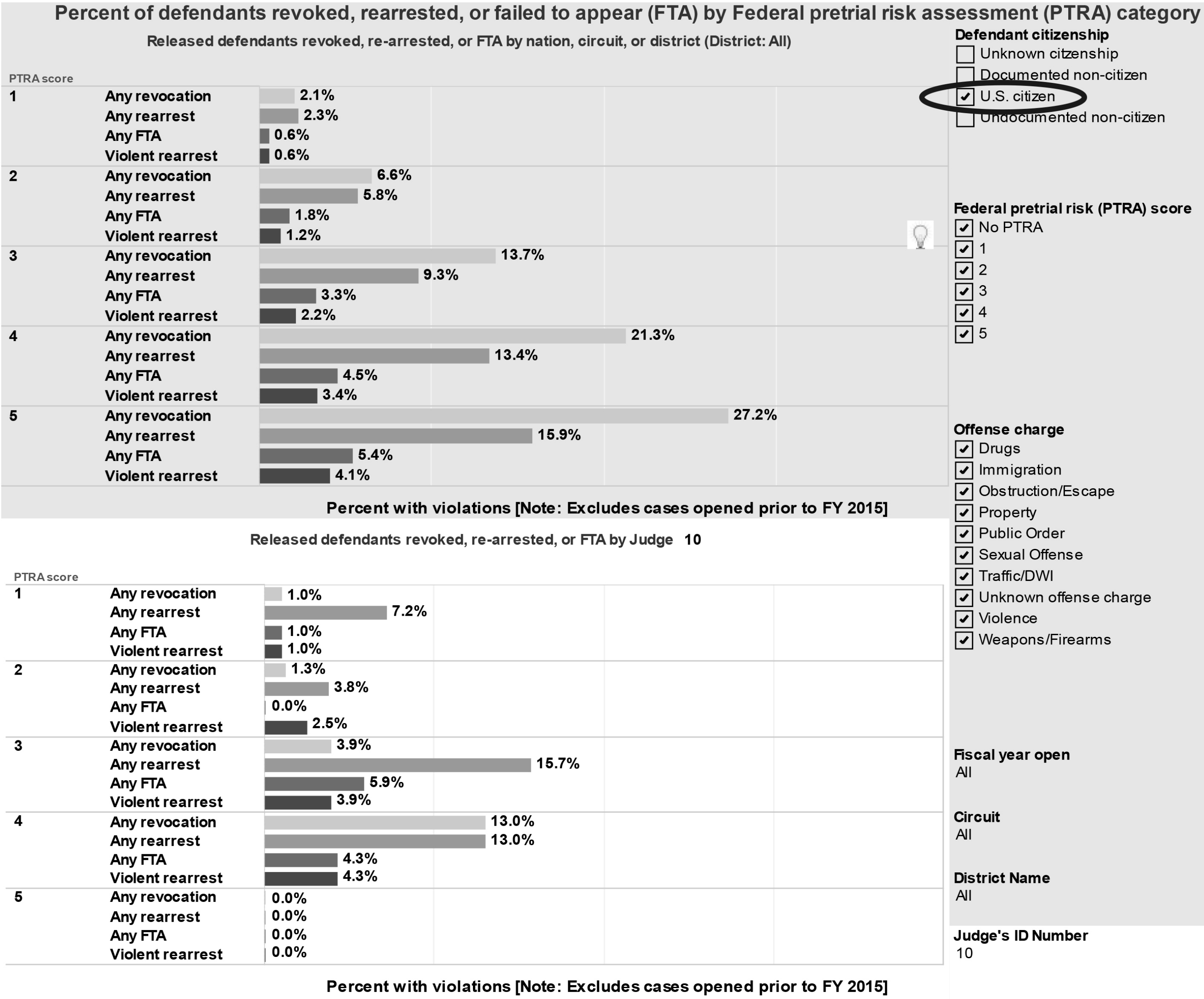


FIGURE 8



Gupta, A., Hansman, C., & Frenchman, E. (2016). The heavy costs of high bail: Evidence from judge randomization. *The Journal of Legal Studies, 45*(2), 471–505.

Heaton, P., Mayson, S. G., & Stevenson, M. (2017). The downstream consequences of misdemeanor pretrial detention. *SSRN, 69*(3), 711–794.

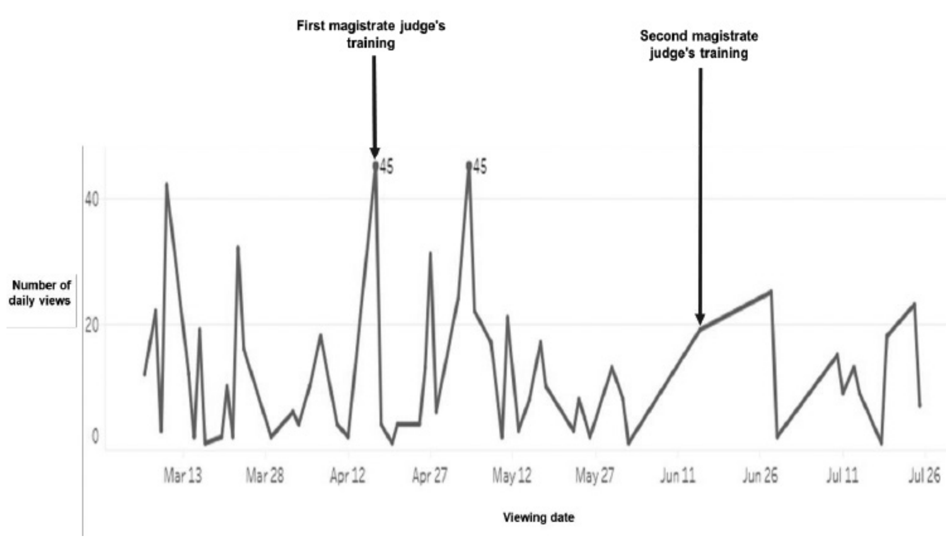
Oleson, J. C., VanNostrand, M., Lowenkamp, C. T., & Cadigan, T. P. (2014). Pretrial detention choices and federal sentencing. *Fed. Probation*, *78*, 12.

Statutes

18 U.S.C. § 3142. Release or detention of a defendant pending trial.

18 U.S.C. § 3154(9). Functions and powers relating to pretrial services.

FIGURE 9



Exploring Probation and Parole Records Using Natural Language Processing: A Case Study of Supervisory Condition Notes

Hadeel Elyazori
Teneshia Thurman
Kevin Lybarger
Faye S. Taxman
George Mason University

PROBATION AND PAROLE conditions are generally set by the Judiciary and/or parole board and define obligations that individuals under supervision must address. Officers typically manage compliance with these conditions. Condition management is an important part of client supervision and requires officers to document various degrees of progress towards meeting these conditions. The documentation of conditions is complicated given the high number of conditions (~8-30) per individual on supervision. Further, the documentation technology is cumbersome, with conditions documented through categorical codes, open-ended text, or a combination of both. This combination of categorical data and unstructured text data complicates large-scale analyses to identify patterns or trends. Consequently, an agency is unlikely to use the text information to review benchmarks or assess the performance of the probation or parole system. Agencies often search for ways to use this textual information, especially since officers are asked or required to enter the data into their automated case management system. The following case study illustrates some natural language processing (NLP) methods that can abstract and summarize the text data and demonstrate the utility of this approach.

NLP is a subfield of artificial intelligence (AI) focused on transforming and interpreting human-generated language. Contemporary AI and NLP are based on machine learning techniques, in which algorithms automatically learn patterns from large data sets. Lauriola et al. provide an overview of NLP, including deep learning techniques.[1] Here we explore the use of NLP-based information extraction techniques, which automatically map unstructured text to a structured semantic representation to facilitate large-scale and real-time analyses. Combining extracted information from officer case notes with the available structured data can create a more holistic understanding of clients and provide actionable insights regarding criminal history, behavior patterns, probation compliance, and other outcomes. A review of the published literature suggests a notable gap in the application of machine learning techniques for information extraction specifically within the context of probation and parole case notes. However, information extraction is well established in other contexts, such as legal documents,[2,3] healthcare,[4] and finance.[5]

The goal of this study is to enable data-centric strategies for better understanding probation and parole practices. We explore officer case notes describing conditions of supervision and use information extraction techniques to convert the unstructured case notes to a semantic representation. We developed a fined-grained, hierarchical annotation (coding) schema for 66 *Condition Category* labels associated with supervisory conditions related to substance use, mental health, treatment programs, community service, education, employment, fines, fees, and other conditions. The 66 *Condition Categories* are related to 10 higher level *Parent Categories*. We annotated the records of over 3,000 clients in a state department of parole and probation and used this annotated corpus to develop information extraction models based on traditional machine learning algorithms

Glossary

- *AI - Artificial Intelligence*
- *BERT - Bidirectional Encoder Representations from Transformers*
- *FLAN - Fine-tuned LAnguage Net*
- *LLM - Large Language Model*
- *NLP - Natural Language Processing*
- *PII - Personally Identifiable Information*
- *RF - Random Forest*
- *SVM - Support Vector Machines*
- *T5 - Text-to-Text Transfer Transformer*
- *TF-IDF - Term Frequency-Inverse Document Frequency*

and state-of-the-art Large Language Models (LLMs). Our results demonstrate the feasibility of using information extraction techniques on probation and parole case notes and provide a foundation for enhancing data analytics within criminal justice settings.

Related Work

AI is increasingly explored within criminal justice, including crime detection,[6] prevention,[7] and forecasting[8,9] and decision support.[10] As examples, Shah et al. used computer vision to forecast crime in videos,[7] and Tollenaar et al. developed machine learning models to predict recidivism risk.[11] Advancements in deep learning (neural networks) are expanding the capabilities and performance of AI in criminal justice and other settings. For example, deep learning crime prediction models can successfully leverage diverse data, including videos, images, audio recordings, and text data, and achieve improved performance over traditional machine learning methods.[8] (See Figure 1.)

Information extraction research within the criminal justice domain has been primarily limited to online law enforcement investigations[12] and legal documents, focusing on names, regulations, legal norms, etc.[2,3] Information extraction research is sparse or non-existent within parole and probation settings. Some research explores parole hearing transcripts, focusing on extracting offenses, gang programming, employment, education, and risk scores.[13] Current literature reviews indicate a scarcity of published research exploring the application of information extraction techniques to parole and probation case notes to understand the supervision process. This lack of published research constitutes a missed opportunity to use technology to improve the supervision and management of offenders. While there is an absence of information extraction work focused on parole and probation case notes, there is a robust body of clinical information extraction research focused on clinician-generated notes describing patients within electronic health records.[4] Clinical data is similar to probation and parole data in that both: i) include structured data and narrative text, ii) contain personally identifying information (PII), iii) document individuals through various domain-specific events, and iv) capture information related to socioeconomic status and health. Our experimentation is informed by clinical information extraction methods.

Information extraction has evolved over time, presenting a continuum from rule-based systems to machine learning and deep learning,[2,3,4,5] where the peformance and capabilities of algorithims have increased over time. Rule-based systems consist of manually curated rules to identify predefined linguistic patterns. Frequently employed traditional machine learning models include logistic regression, Random Forest (RF), and Support Vector Machines (SVM).[2,3] RF ensembles multiple decision trees to make predictions (see Figure 2A), and SVM finds the optimal boundary to separate categories[2,3] (see Figure 2B). For traditional methods, a common approach for converting text to input features is Term Frequency-Inverse Document Frequency (TF-IDF), which assigns weights to words based on their frequency[3] (see Figure 2C). TF-IDF word weighting assigns higher values to words that are more frequent in a document and less frequent in the other documents in the corpus. More recently, neural networks, like Convolutional Neural Networks and Recurrent Neural Networks, have achieved prominence over traditional methods due to their capacity for automated feature learning and ability to model complex relationships within text data.[2,3,4,5]

FIGURE 1

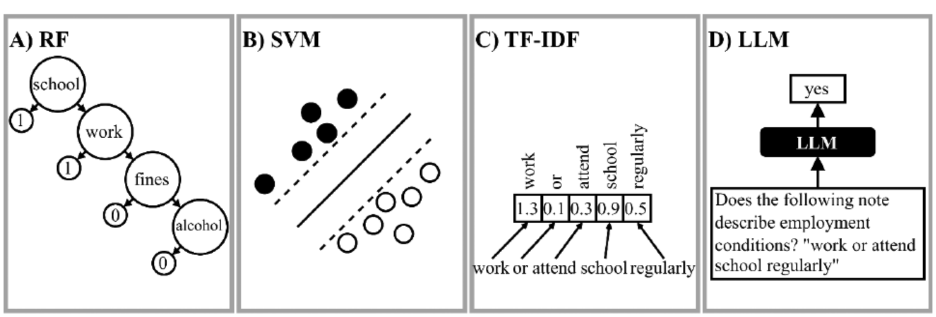


Figure 1. Modeling architectures. A) Random forest (RF) – presents a single example decision tree; B) Support Vector Machines – illustrates defining a decision boundary that optimally separates samples; C) Term frequency-inverse document frequency (TF-IDF) – presents example mapping of input text to a feature vector; and D) Large language model (LLM) – presents an example where the model input and output are text.

LLMs, like ChatGPT,[14] currently dominate the NLP landscape and achieve state-of-the-art performance in myriad tasks, including information extraction. LLMs are built on transformer architectures and include millions to trillions of trainable parameters. The typical training approach involves *pre-training* on extensive unlabeled text corpora to acquire a generalized understanding of language, followed by *fine-tuning* (supervised learning) on labeled data to learn a specific task. This transfer learning paradigm is particularly advantageous in domains where annotated data is limited, a condition relevant to corrections and community supervision settings. To address privacy concerns related to PII, we focus on two publicly available architectures: Bidirectional Encoder Representations from Transformers (BERT)[15] and Text-to-Text Transfer Transformer (T5)[16] (see Figure 2D). BERT encodes text by transforming input word sequences into vectors that can be used for classification. BERT has achieved state-of-the-art performance in many information extraction tasks across domains.[2,3,4,5] T5 is a generative model that transforms input text to output text and can be used for many tasks, including classification. T5 has achieved state-of-the-art performance in many tasks, including the extraction of social determinants of health in clinical notes.[17]

Methods and Materials

Data

In this study, we used client case plan data from a parole and probation agency located in a mid-Atlantic state. The data includes over 3,000 unique clients, which covers cases opened from 2017-2021. Client case plans describe the requirements and conditions an individual must follow during supervision. Probation/parole officers use an agency's database to document conditions and design goals to achieve them. The goals can refer to activities such as random urinalysis, taking prescribed medication, obtaining mental health evaluation, participating in mental health treatment, and other requirements. In our study, the agency-provided data included 120 *Condition Codes*, each with a corresponding *Condition Description* that is constant across all records. For example, agency-provided *Condition Codes* 9532 and 16028 have the *Condition Descriptions* "Other" and "Additional Drug condition," respectively. Within the dataset, there were 34 *Condition Codes* that also included an officer-generated *Condition Note* documenting case plan details through unstructured narrative text. For example, the *Condition Code* 9532 with *Condition Description* "Other" serves as one of several catchall codes for conditions that do not easily fit more specific codes. The officer-generated *Condition Notes* for *Condition Code* 9532 document a wide range of conditions, such as "Defendant not to drive," "Seek employment/school," or "Gun registry." The 34 *Condition Codes* with associated *Condition Notes* include: 1) *Other* – 12 codes were described as "other" and serve as a catchall for undefined conditions; 2) *Programs* – 4 codes require officers to specify particular programs, for example behavioral health, domestic violence, veteran, family counseling, and vocational programs; 3) *Substance Use* – 5 codes pertain to drug or alcohol conditions; 4) *Victim* – 2 codes were victim-focused conditions; 5) *Sex Offender* – 2 codes were related to sex offenders' special conditions; and 6) *Additional Conditions* – 9 codes addressed specific requirements or restrictions, which involved completion of assigned tasks or community service, financial obligations such as court costs and restitution, geographical limitations, and specified durations of home confinement or other monitoring requirements. Officers can amend their case plans through supervision. Each client may have multiple parole or probation cases, and each case can include multiple conditions. We treat each condition record (*Condition Code*, *Condition Description*, and *Condition Note*) as a sample or record. In addition to conditions, the data set includes: 1) *case type* – parole vs. probation and 2) *case level* – low, low-moderate, moderate, maximum, special cases, or violent.

Annotation

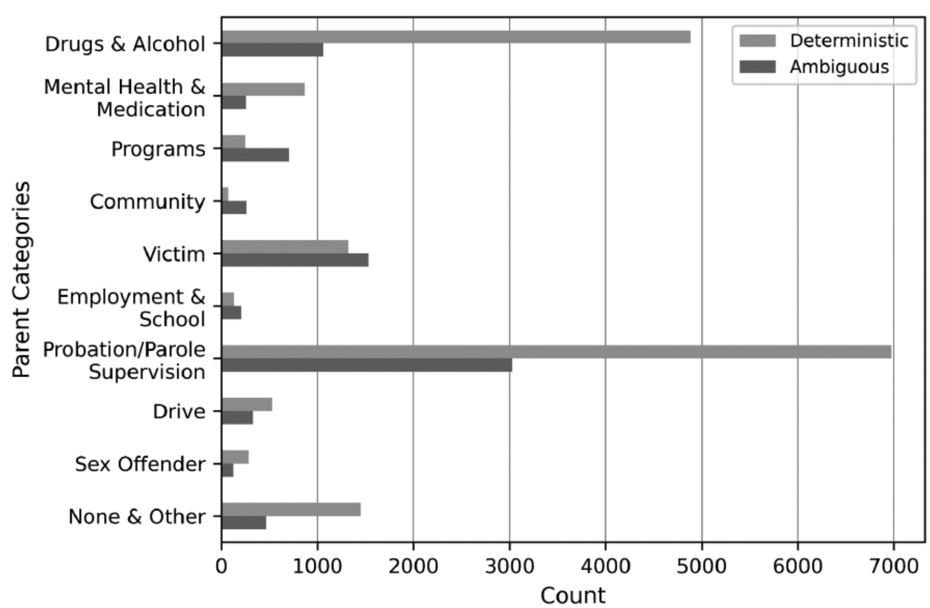
The primary objective of the annotation was to identify and categorize the 34 *Condition Codes* that included an officer-generated *Condition Note*; however, we developed a comprehensive set of *Condition Category* labels that summarized the meaning of all 120 *Condition Codes*. Officers manually type the *Condition Notes*, requiring a comprehensive review and categorization process. Based on our review of the data, we developed a set of 66 *Condition Category* labels to map the condition records, including unstructured *Condition Note* information, to a fixed set of classes. Annotation involved assigning one or more of the researcher-defined 66 *Condition Category* labels to the agency-provided records. Table 1 summarizes

TABLE 1
Condition Category Hierarchy

Parent Categories	Condition Categories	
Drugs & Alcohol	• Drug/Alcohol Testing • No Alcohol • No Drugs • No Drugs or Alcohol • No Specific Substance • Misc. Substance Abuse Cond. • Undergo Drug/Alcohol Eval. • Attend Alcohol Tx • Attend Drug Tx	• Attend Substance Use Prog. • Drug-Related Cond. • Alcohol Cond. • Attend Alcohol Prog. • Attend Drug Prog. • Undergo Alcohol Screening & Tx • Undergo Drug Screening & Tx • Other Alcohol Related Cond. • Other Drug Related Cond.
Mental Health & Medication	• Mental Health Eval. • Ordered to Take Medication • Psychiatric Cond. • Different Cond. Eval.	• Attend Counseling Aftercare • Attend Mental Health Court • Attend Outpatient Tx Prog.
Programs	• Participate in Self-Help Group • Undergo Anger Management • Misc. or Unknown Prog. Cond. • Reentry Prog. Cond. • Attend Prog. for Veterans	• Attend a Behavioral Health Prog. • Attend Parenting Prog. • Attend Drug Court • Supervised by Mental Health Agent/Unit • None specified Tx
Community	• Community Service Cond.	• Reentry Into Community Cond.
Victim	• Victim-Related Cond. • Other Victim Related Cond.	• Attend Victim Prog.
Employment & School	• Employment or School Cond.	• Employment or School Prog.
Probation/Parole Supervision	• Curfew Cond. • Gen. Probation & Parole Cond. • Undergo Record Check • Provide DNA • Appear in Court Cond. • Pay Fines/Fees Cond. • Movement Restriction • Supervision Relocation	• Fines & Fees Waived • Restitution Cond. • Gun-Related Cond. • Allowed to Leave the State • Non-standard Probation Cond. • Participate in Probation/Parole Prog. • Apology Letter Requirement
Drive	• Driving/Driver's License	
Sex Offender	• Sex Offender Cond.	
None & Other	• No Special Cond. • Other • Undergo an Eval. • COVID-19 Cond.	• Family-Related Cond. • Forfeit Items • Housing Cond.

Abbreviations: condition (cond.), evaluation (eval.), general (gen.), miscellaneous (misc.), program (prog.), and treatment (Tx).

FIGURE 2
Distribution of Condition Categories



Deterministic indicates the Condition Category label can be assigned based solely on the agency-provided Condition Code. Ambiguous indicates the Condition Category label assignment requires interpretation of the office-generated Condition Note text.

the assigned labels, which are hierarchically arranged with 66 *Condition Category* labels assigned to 10 *Parent Categories*. Among the 120 *Condition Codes*, 86 *Condition Codes* do not include *Condition Notes* and always correspond with the same *Condition Category*, so they can be deterministically assigned a *Condition Category* label; and 34 *Condition Codes* include *Condition Notes* that must be interpreted to resolve ambiguity regarding the relevant *Condition Category* label. Before manual coding of the case plan requirements, *Condition Categories* were automatically assigned to the 86 deterministic *Condition Codes* that do not include associated *Condition Notes*, and manual annotation focused on resolving the ambiguity associated with the 34 *Condition Codes* that included narrative text through *Condition Notes*. During the annotation process, new *Condition Category* labels were added to the label set if the condition did not align with existing categories. Samples were annotated by three individuals with domain expertise, including backgrounds in criminology and policy. Extensive annotation training ensured data quality and annotation consistency.

Figure 2 summarizes the distribution of the *Parent Category* labels broken down by: 1) *deterministic* – the record does not include officer-generated text (*Condition Note*), and the *Condition Category* label can be assigned to the record based solely on the *Condition Code* and 2) *ambiguous* – officer-generated *Condition Note* text must be interpreted to determine the appropriate category label. In total, 48 percent of records require interpretation of the officer-generated *Condition Note*, indicating the text's importance in understanding the assigned condition.

Condition Category Dependence

To better understand the relationship between the *Condition Category* labels and the client case type and level, we performed Chi-squared test of independence between each *Condition Category* label and the case type and level. The case type is binary (probation vs. parole). The case level is multiclass, and we converted the case level labels to a binary one-versus-rest representation before performing the statistical test.

Information Extraction

We explored the *Condition Category* prediction task for the records with a *Condition Note* (ambiguous records in Figure 2), using traditional machine learning models and LLMs. For all experiments, the model input is the client record (*Condition Code*, *Condition Description*, and *Condition Note*). In our annotation scheme, each record can be assigned multiple *Condition Category* labels, so we treat this task as a multi-label binary prediction task, where each record is assigned a set of 66 binary labels (1 indicates category relevant, and 0 indicates category irrelevant). Figure 3 presents an overview of the modeling approaches, including examples of how the record is represented in the input. The *Condition Code* and *Condition Description* are included with the *Condition Note* in the model input to provide important context for interpretation.

Traditional Machine Learning

We explored two traditional machine learning models: 1) RF and 2) SVM. The input to these models includes the *Condition Code* and TF-IDF representation of the *Condition Description* and *Condition Note*. The RF/SVM models learn feature weights for the features to predict the *Condition Category* labels. Separate RF and SVM models were developed for each *Condition Category*, and predictions from the category-specific models were combined to form a set of predictions for each record. Figure 3A presents an example of a single RF/SVM classifier, where the output is a binary prediction for a single *Condition Category*.

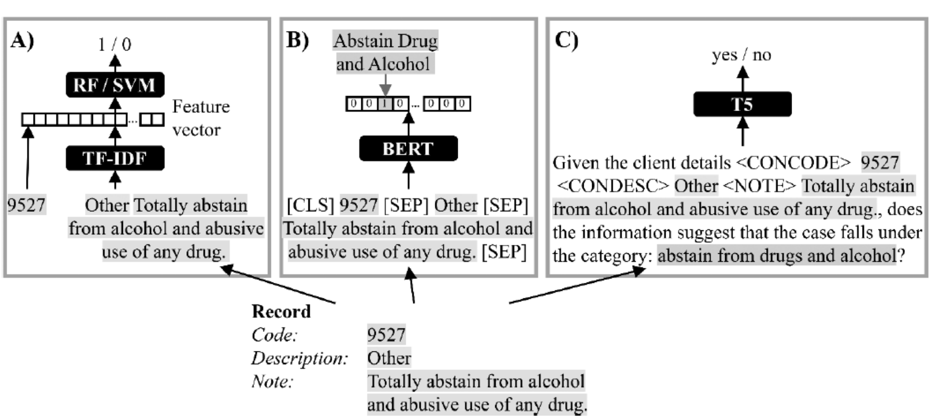
LLMs

We explored two LLMs: BERT and T5. BERT is pretrained on a large body of text to learn a general representation of language. In this pretraining, special tokens are included to define the input format, including: CLS – specifies the start of the input and SEP – serves as a separator for different inputs. As shown in Figure 3B, the BERT input consists of the *Condition Code*, *Condition Description*, and *Condition Note* separated by the SEP token. BERT maps this input text to an output vector, and separates linear functions for each *Condition Category* to generate binary predictions. In this configuration, a single BERT model can generate all 66 multi-label predictions. As is common practice, we started with a pretrained BERT model, then trained the BERT model and output linear functions on the labeled data. As presented in Figure 3C, we used T5 to assign *Condition Category* labels using a question-answering (QA) setting. In this QA setting, a separate yes/no question is formulated for each *Condition Category*, and the set of yes/no questions spanning all *Condition Category* labels is applied to each record. The input to T5 includes a *Condition Category*-specific question and the *Condition Code*, *Condition Description*, and *Condition Note* separated by special tokens (e.g., <Code> or <Description>) to differentiate input information. The T5 output is a "yes" / "no" answer to the *Condition Category*-specific question.

Experimental Paradigm

Modeling was implemented using the Python packages Scikit-learn[18] and Transformers.[19] Records were divided into three subsets at the client level: 70 percent training, 10 percent validation, and 20 percent testing. The optimal configuration (hyperparameters) for each model was determined by training models on the training set and evaluating performance on the validation set. We report the performance on the withheld test set using the optimal configurations. Detailed model configurations are presented in the Appendix.

FIGURE 3
Information Extraction Architectures



Performance

Performance is evaluated using precision, recall, and F1, as defined in Equation 1. Given the high number of *Condition Category* labels, we report the micro-averaged performance at the *Parent Category* level and include individual *Condition Category* performance in the Appendix. The statistical significance of the results was evaluated using a pairwise non-parametric test (bootstrap test, p-value<0.05).[20]

Results

Condition Category Dependence

Our study first focused on comparing our *Condition Category* labels with the agency-provided case types and case levels through Chi-squared tests of independence summarized in Table 2. For space, Table 2 only presents 32 of the 66 *Condition Category* labels that are dependent on case type or level. The triangles (▲ or ▼) indicate that the *Condition Category* label and case level or type (the *variables*) are dependent. An upward-facing triangle (▲) indicates the variables co-occur more frequently and a downward-facing triangle (▼) indicates the variables co-occur less frequently than expected, if the variables were independent. The diversity in the conditions across different case levels and types illustrates the complexity of decision-making and the tailored strategies employed to address the varying needs and risks associated with each case; however, several themes emerged from this analysis. Probation tends to have higher rates than parole for conditions related to drugs and alcohol, self-help, anger management, community service, victims, waiving fees, guns, driver's licenses, and forfeiture of items. Conversely, parole has higher rates than probation for conditions related to employment, curfew, paying fees, and sex offender conditions. Lower level offenders (low and low/moderate) tend to have higher prevalence than higher level offenders (moderate, maximum, special case, and violent) for conditions related to drugs and alcohol, self-help, community service, attending victim programs, and driver's license. Conversely, higher level offenders have higher rates than lower level offenders for conditions related to anger management, victim conditions (other than victim programs), curfew, and sex offender conditions.

Classification Performance

Table 3 presents the prediction performance on the withheld test set for the *Condition Category* labels micro-averaged for each *Parent Category*. In information extraction research, performance varies by task and data set, and there are not predefined thresholds for good/acceptable performance; however, we consider performance ≥ 0.90 F1 to be very high. The LLMs (BERT and T5) outperformed the traditional machine learning models (RF and SVM) in the overall performance, as well as the performance in 5 of the 10 *Parent Categories*, with significance, demonstrating the natural language understanding capabilities of the LLMs. Among all models, T5 achieved the highest overall performance and *Mental Health & Medication* performance with significance. These results demonstrate the feasibility of developing high-performing information extraction models for probationary notes and highlight the value of using LLMs. Table 4 in the Appendix presents the performance for the individual *Condition Category* labels.

Error Analysis

Each *Parent Category* includes a set of topically relevant *Condition Categories*. The performance for the *Parent Categories* tends to be higher when there are fewer associated *Condition Categories*, as the classification models need to disambiguate fewer topics. For example, the T5 performance is ≥ 0.97 F1 for the *Parent Categories* – *Community*, *Victim*, and *Drive* – which have 2, 3, and 1 child labels respectively. Additionally, the highest performing *Parent Categories* include *Condition*

EQUATION 1
Precision, Recall, and F1 Formulas

$$Precision = \frac{TP}{TP + FP}; \quad Recall = \frac{TP}{TP + FN}; \quad F1 = 2\frac{Precision \times Recall}{Precision + Recall}$$

Abbreviations: true positive (TP), false positive (FP), and false negative (FN)

TABLE 2
Condition Category Dependency

Parent Categories	Level - Low	Level - Low/moderate	Level - Moderate	Level - Maximum	Level - Special Case	Level - Violent	Type - Probation	Type - Parole	Condition Categories	Level - Low	Level - Low/moderate	Level - Moderate	Level - Maximum	Level - Special Case	Level - Violent	Type - Probation	Type - Parole	Condition Categories
Drugs & Alcohol		▲				▼			Drug/Alcohol Testing	▲				▼	▼			Attend Substance Use Prog.
		▲				▲			Misc. Substance Abuse Cond.	▲	▲	▼	▼		▼	▲	▼	Attend Alcohol Prog.
		▲				▼	▲	▼	Attend Alcohol Tx					▼				Attend Drug Tx
					▼		▲	▼	Undergo Drug/Alcohol Eval.									
Mental Health & Medication				▲	▲				Psychiatric Cond.					▲				Attend Mental Health Court
Programs	▲					▼	▲	▼	Participate in Self-Help Group				▲					Attend a Behavioral Health Prog.
	▼	▼	▲	▲			▲	▼	Undergo Anger Management					▲				Attend Parenting Prog.
						▲			Misc. or Unknown Prog. Cond.				▲					Attend Drug Court
			▲						Reentry Prog. Cond.									
Community	▲				▼	▼	▲	▼	Community Service Cond.									
Victim	▼	▼	▲	▲	▲		▲	▼	Victim-Related Cond.	▲	▲	▼	▼		▼	▲	▼	Attend Victim Prog.
Employment & School	▼		▲				▼	▲	Employment or School Cond.						▲			Employment or School Prog.
Probation/ Parole Supervision				▲	▲	▲	▼	▲	Curfew Cond.				▲			▲	▼	Fines & Fees Waived
				▼			▼	▲	Gen. Probation & Parole Cond.			▲			▲			Restitution Cond.
	▲	▼		▼			▲	▼	Undergo Record Check	▼			▲		▲	▲	▼	Gun-Related Cond.
	▲			▼	▼	▲	▼	▲	Pay Fines/Fees Cond.									
Drive	▲		▲	▼	▼	▼	▲	▼	Driving/Driver's License									
Sex Offender	▼	▼	▲	▲	▲	▼	▼	▲	Sex Offender Cond.									
None & Other		▼		▲		▲	▲	▼	Forfeit Items							▼	▲	COVID-19 Cond.

An upward or downward facing triangle (▲ or ▼) indicates the Condition Category label and case level or type are dependent (p<0.05, null hypothesis of independence rejected). An upward facing triangle (▲) indicates the variables co-occur more frequently than expected if independent, and a downward facing triangle (▼) indicates the variables co-occur less frequently than expected if independent. Abbreviations: condition (cond.), evaluation (eval.), general (gen.), miscellaneous (misc.), program (prog.), and treatment (Tx).

Categories with very consistent linguistic cues (keywords). For example: i) *Community Service Condition* – "community service" or time commitment, like "10 hours per week"; ii) *Victim-Related Condition* – "no contact with" or "do not enter"; iii) *Attend Victim Program* – "victim impact panel" or "VIP"; and iv) *Driving/Driver's License* – "drive," "interlock," or "license."

The performance for *Parent Categories* tends to be lower when there are more associated *Condition Categories*, as the models must distinguish between more closely related topics. For example, the T5 performance is ≤ 0.82 F1 for the *Parent Categories – Drug & Alcohol*, *Programs*, and *None & Other* – which have 18, 10, and 8 child labels respectively. Within these *Parent Categories*, the individual *Condition Category* performance varies, and performance decreases as linguistic diversity increases. For example, the *Condition Category Miscellaneous or Unknown Program Condition* is a catchall for requirements related to a range of programs, and the notes contain diverse language, references to specific treatment facilities, and ambiguous statements like "successfully complete treatment." As another example, the *Condition Category – Attend Substance Use Program* – includes notes describing several different specific treatment programs and facilities and includes less common shorthand, like "ALC PGM" for "Alcohol Program."

Discussion

The overarching goal of this study is to enable probation and parole agencies to use the information captured in officer-generated notes in large-scale and real-time analyses. This goal is highly significant, due to the prevalence of open-ended text fields in management information systems, importance of the textual information, and challenges associated with converting this textual information into quantifiable data. Through NLP information extraction techniques, the unstructured text can be converted to a structured representation to examine patterns and assess performance at all levels, including the program, officer, and individual under supervision. Agencies currently grapple with the complexity of summarizing these text data, but the strategies presented in this case study demonstrate how NLP can generate usable metrics that can easily be combined with existing categorial data. While these strategies require specific technical expertise, this work illustrates the value of AI methods.

In our study, the LLMs (BERT and T5) outperformed traditional machine learning models (RF and SVM). For the traditional models, all model learning originates from annotated training data. However, the LLMs use transfer learning, where the models first pretrain on large corpora of unlabeled text to learn language understanding and then fine-tune (train) on the annotated training data to learn the target task. The improved performance of the LLM can be attributed to the success of this learning transfer, which provides a general understanding of language. The improved performance of the T5 model relative to BERT can be attributed to the larger model size (higher number of parameters) and larger pretraining corpus.

We are unaware of any prior information extraction work exploring officers' documentation of parole and probation conditions. Our results demonstrate the feasibility of using information extraction techniques in this setting by achieving high performance across most of the *Parent Categories*. The use of NLP with correctional system data, including parole and probation notes, has the potential to improve management and supervision by enabling the automatic analysis of vast amounts of information-dense text data. It can provide a richer, data-driven understanding of offender behavior and risks and could lead to more tailored intervention strategies and more informed decision-making processes, ultimately contributing to improved rehabilitation and public safety.

This research has key limitations related to data heterogeneity. First, we explored a moderately sized client population from a single agency, and the populations in the analyzed data set may not be representative of other agencies. The conditions and documentation practices, including the authoring of notes by officers, may vary by agency, and additional work is needed to understand the variability of the conditions and notes across institutions. Second, we explored officer descriptions of conditions, which represent only one of many types of free-text records within correctional data. Additional analyses with more comprehensive text record types are needed to understand the feasibility and challenges associated with applying information extraction techniques more broadly within correctional system data.

Conclusions

We explored a corpus of officer-generated notes documenting the parole and probation conditions of clients under supervision and investigate the use of state-of-the-art information extraction techniques. We annotated the records of over 3,000 clients with a fine-grained annotation schema of 66 *Condition Categories* and developed information extraction models based on traditional machine learning methods and LLMs. The LLMs outperformed the traditional machine learning methods, with the generative T5 model achieving the best overall performance at 0.89 F1. This high performance demonstrates the feasibility of using NLP in this parole and probation setting and provides a foundation for future exploration of correctional system data.

Ethics

We had the necessary approvals from our institution's Institutional Review Board (IRB) to obtain, store, and analyze the probation and parole data set. All researchers and annotators received the necessary human subjects

TABLE 3
Classification Results on Withheld Test Set*

Parent Category	# Human-annotated Labels	F1 Micro Average			
		RF	SVM	BERT	T5
Drugs & Alcohol	219	0.61	0.71	0.79*	**0.82***
Mental Health & Medication	46	0.68	0.74	0.78*	**0.86*†**
Programs	122	0.66	0.70	**0.76**	0.75
Community	20	0.80	0.86	0.97*	**0.98***
Victim	194	0.94	0.94	0.96	**0.97**
Employment & School	41	0.85	0.89	0.92	**0.93**
Probation/Parole Supervision	347	0.74	0.78	0.86*	**0.90***
Drive	66	0.98	0.98	0.97	**0.99**
Sex Offender	25	0.89	0.92	0.94*	**0.99***
None & Other	71	0.70	0.71	**0.73**	**0.73**
OVERALL MICRO AVG.	1151	0.77	0.81	0.85*	**0.89*†**

** Indicates LLM significantly outperforms traditional model (RF and SVM). † Indicates T5 significantly outperforms BERT.*

training to interact with the client data, including the PII.

Acknowledgments

We would like to thank the state agency for offering this data and giving us the opportunity to understand how to use various technologies to explore strategies to abstract information. We acknowledge that Dr. Faye S Taxman is the PI of this project and supported this study.

References

1. I. Lauriola, A. Lavelli and F. Aiolli, "An introduction to deep learning in natural language processing: Models, techniques, and tools," *Neurocomputing,* vol. 470, pp. 443-456, 1 2022.
2. F. Solihin, I. Budi, R. F. Aji and E. Makarim, "Advancement of information extraction use in legal documents," *International Review of Law, Computers and Technology,* vol. 35, no. 3, pp. 322-351, 2021.
3. C. Sansone and G. Sperlí, "Legal information retrieval systems: State-of-the-art and open issues," *Information Systems,* vol. 106, 5 2022.
4. Y. Wang, L. Wang, M. Rastegar-Mojarad, S. Moon, F. Shen, N. Afzal, S. Liu, Y. Zeng, S. Mehrabi, S. Sohn and H. Liu, "Clinical information extraction applications: A literature review," *Journal of Biomedical Informatics,* vol. 77, pp. 34-49, 1 2018.
5. M. H. A. Abdullah, N. Aziz, S. J. Abdulkadir, H. S. A. Alhussian and N. Talpur, "Systematic literature review of information extraction from textual data: Recent methods, applications, trends, and challenges," *IEEE Access,* vol. 11, pp. 10535-10562, 2023.
6. C. Rigano, "Using artificial intelligence to address criminal justice needs," *National Institute of Justice Journal,* vol. 280, pp. 1-10, 2019.
7. N. Shah, N. Bhagat and M. Shah, "Crime forecasting: A machine learning and computer vision approach to crime prediction and prevention," *Visual Computing for Industry, Biomedicine, and Art,* vol. 4, no. 1, p. 9, 2021.
8. V. Mandalapu, L. Elluri, P. Vyas and N. Roy, "Crime prediction using machine learning and deep learning: A systematic review and future directions," *IEEE Access,* vol. 11, pp. 60153-60170, 2023.
9. O. Kounadi, A. Ristea, A. Araujo Jr and M. Leitner, "A systematic review on spatial crime forecasting," *Crime Science,* vol. 9, no. 1, p. 7, 5 2020.
10. J. Mitchell, S. Mitchell and C. Mitchell, "Machine learning for determining accurate outcomes in criminal trials," *Law, Probability and Risk,* vol. 19, no. 1, p. 43–65, 2020.
11. N. Tollenaar and P. G.M. van der Heijden, "Which method predicts recidivism best?: A comparison of statistical, machine learning and data mining predictive models," *Journal of the Royal Statistical Society Series A: Statistics in Society,* vol. 176, no. 2, p. 565–584, 2013.
12. M. Edwards, A. Rashid and P. Rayson, "A systematic survey of online data mining technology intended for law enforcement," *ACM Computing Surveys,* vol. 48, no. 1, 9 2015.
13. J. Hong, C. Voss and C. D. Manning, "Challenges for information extraction from dialogue in criminal law," in *Workshop on NLP for Positive Impact*, 2021.
14. OpenAI, "GPT-4 technical report," 3 2023.
15. J. Devlin, M.-W. Chang, K. Lee and . K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in *Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, Minneapolis, MN, 2019.
16. C. Raffel, N. Shazeer, A. Roberts, K. Lee, S. Narang, M. Matena, Y. Zhou, W. Li and P. J. Liu, "Exploring the limits of transfer learning with a unified text-to-text transformer," *Journal of Machine Learning Research,* vol. 21, pp. 1-67, 2020.
17. B. Romanowski, A. Ben Abacha and Y. Fan, "Extracting social determinants of health from clinical note text with classification and sequence-to-sequence approaches," *Journal of the American Medical Informatics Association : JAMIA,* vol. 30, no. 8, pp. 1448-1455, 7 2023.
18. F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. P. Passos, D. Cournapeau, M. Brucher, M. Perrot and É. Duchesnay, "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research,* vol. 12, no. 85, pp. 2825-2830, 2011.
19. T. Wolf, L. Debut, V. Sanh, J. Chaumond, C. Delangue, A. Moi, P. Cistac, T. Rault, R. Louf, M. Funtowicz, J. Davison, S. Shleifer, P. Von Platen, C. Ma, Y. Jernite, J. Plu, C. Xu, T. Le Scao, S. Gugger, M. Drame, Q. Lhoest and A. M. Rush, "Transformers: State-of-the-Art Natural Language Processing," in *Association for Computational Linguistics*, 2020.
20. T. Berg-Kirkpatrick, D. Burkett and D. Klein, "An Empirical Investigation of Statistical Significance in NLP," in *Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, Jeju Island, Korea, 2012.

Appendix

Model Configuration

Each architecture includes some model-specific configuration. For the RF, the optimum hyperparameters include class weight = balanced subsample, maximum depth = 50, and number of estimators = 200. For the SVM, the optimum hyperparameters include C = 100. For BERT, we started with the pretrained model *bert-base-uncased* and trained the model for 29 epochs. For T5, we started with the pretrained *model flan-t5-large* and trained the model for 20 epochs.

TABLE 4
Detailed Performance for the Individual Condition Category Labels

Parent Category	Condition Category	# Human-annotated Labels	F1 Micro Average RF	SVM	BERT	T5
Drugs & Alcohol	Drug/Alcohol Testing	34	0.81	0.87	0.86	0.90
	No Alcohol	4	0.00	0.00	0.00	0.03
	No Drugs	2	0.00	0.50	0.00	0.11
	No Drugs or Alcohol	22	0.86	0.93	0.82	0.83
	No Specific Substance	0	0.00	0.00	0.00	0.00
	Misc. Substance Abuse Cond.	0	0.00	0.00	0.00	0.00
	Undergo Drug/Alcohol Evaluation	16	0.40	0.46	0.73	0.69
	Attend Alcohol Tx	10	0.15	0.78	0.83	0.84
	Attend Drug Tx	12	0.29	0.50	0.80	0.44
	Attend Substance Use Program	79	0.58	0.64	0.79	0.68
	Drug-Related Cond.	0	0.00	0.00	0.00	0.00
	Alcohol Cond.	1	0.00	0.00	0.00	0.02
	Attend Alcohol Program	33	0.71	0.82	0.86	0.71
	Attend Drug Program	5	0.00	0.60	0.75	0.45
	Undergo Alcohol Screening & Tx	0	0.00	0.00	0.00	0.00
	Undergo Drug Screening & Tx	1	0.00	0.00	0.00	0.02
	Other Alcohol Related Cond.	0	0.00	0.00	0.00	0.00
	Other Drug Related Cond.	0	0.00	0.00	0.00	0.00
Mental Health & Medication	Mental Health Evaluation	2	1.00	1.00	1.00	1.00
	Ordered to Take Medication	7	0.50	0.50	0.60	0.76
	Psychiatric Cond.	31	0.76	0.77	0.84	0.89
	Different Cond. Evaluation	0	0.00	0.00	0.00	0.00
	Attend Counseling Aftercare	6	0.00	0.67	0.55	0.71
	Attend Mental Health Court	0	0.00	0.00	0.00	0.00
	Attend Outpatient Tx Program	0	0.00	0.00	0.00	0.00
Programs	Participate in a Self-Help Group	31	0.87	0.83	0.92	0.85
	Undergo Anger Management	32	0.91	0.90	0.91	0.92
	Misc. or Unknown Program Cond.	34	0.11	0.44	0.55	0.59
	Reentry Program Cond.	7	0.73	0.92	1.00	1.00
	Attend Program for Veterans	0	0.00	0.00	0.00	0.00
	Attend a Behavioral Health Program	8	0.50	0.33	0.50	0.52
	Attend Parenting Program	3	0.50	0.50	1.00	1.00
	Attend Drug Court	4	0.40	0.40	0.86	0.82
	Supervised by Mental Health Agent/Unit	0	0.00	0.00	0.00	0.00
	None specified Tx	3	0.00	0.00	0.00	0.00
Community	Community Service Cond.	20	0.80	0.86	0.97	0.98
	Reentry Into Community Cond.	0	0.00	0.00	0.00	0.00
Victim	Victim-Related Cond.	157	0.93	0.93	0.96	0.97
	Other Victim Related Cond.	0	0.00	0.00	0.00	0.00
	Attend Victim Program	37	0.97	1.00	0.96	0.98
Employment & School	Employment or School Cond.	36	0.91	0.93	0.97	0.97
	Employment or School Program	5	0.00	0.57	0.55	0.64
Probation/Parole Supervision	Curfew Cond.	10	0.17	0.57	0.70	0.54
	General Probation & Parole Cond.	17	0.71	0.65	0.62	0.59
	Undergo Record Check	11	0.90	1.00	1.00	1.00
	Provide DNA	1	0.00	0.00	1.00	0.39
	Appear in Court Cond.	9	0.33	0.57	0.88	0.75
	Pay Fines/Fees Cond.	143	0.92	0.92	0.96	0.97
	Movement Restriction	4	0.75	0.89	0.80	0.96
	Supervision Relocation	5	0.00	0.50	0.60	0.30
	Fines & Fees Waived	8	0.77	0.67	0.93	0.95
	Restitution Cond.	31	0.65	0.75	0.91	0.68
	Gun-Related Cond.	30	0.76	0.82	0.98	0.99
	Allowed to Leave the State	15	0.55	0.67	0.76	0.80
	Non-standard Probation Cond.	59	0.32	0.46	0.62	0.37
	Participate in Probation/Parole Program	3	1.00	1.00	0.86	1.00
	Apology Letter Requirement	1	0.00	0.00	0.67	0.72
Drive	Driving/Driver's License	66	0.98	0.98	0.97	0.99
Sex Offender	Sex Offender Cond.	25	0.89	0.92	0.94	0.99
None & Other	No Special Cond.	1	0.00	0.00	0.00	0.00
	Other	29	0.40	0.43	0.44	0.50
	Undergo an Evaluation	4	0.40	0.33	0.67	0.67
	COVID-19 Cond.	11	0.95	0.95	1.00	1.00
	Family-Related Cond.	2	0.00	0.00	0.00	0.00
	Forfeit Items	24	0.96	0.96	0.98	1.00
	Housing Cond.	0	0.00	0.00	0.00	0.00
	OVERALL MICRO AVG.	**1151**	**0.77**	**0.81**	**0.85**	**0.89**

Automated Extraction of Substance Use and Co-occurring Disorders from Probation Records[1]

Karine Megerdoomian, The MITRE Corporation
Charles E. Horowitz, The MITRE Corporation
Amy B. Marsh, The MITRE Corporation

INDIVIDUALS IN THE criminal justice system who have a history of substance use disorder (SUD) have been shown to display a higher rate of rearrest and recidivism (Stahler et al., 2013; Fazel et al., 2016). Of particular concern are cases where the substance use disorder appears with co-occurring disorders[1] (CODs); these individuals are less likely to enter and successfully complete treatment and are at an even greater risk for criminal relapse (Monahan, 1992; Drake & Wallach, 1989). Comorbidity of mental illness with addiction is also associated with numerous negative health outcomes, increased risk of homelessness, loss of employment, or self-harm (SAMHSA, 2022). Given their multiple needs, these individuals often require specialized interventions, providing integrated mental health and substance use services, to facilitate their reintegration within society. Comprehensive research on the prevalence, trends, and correlates of SUDs and co-occurring mental illnesses is necessary to guide evidence-based, timely, and effective policies and programs aimed at increasing public safety and reducing recidivism (Fearn et al., 2017). Yet, there is a dearth of research on this topic within the justice system, and most studies on this population are limited to small-scale cohort studies.

The U.S. Probation and Pretrial Services Office (PPSO) produces billions of pages of information on individuals under supervision, including detailed social and psychological history on prior and current substance use, official diagnoses, and treatment information. This information is critical for probation officers and district chiefs to better assist the substance use population and to guide efficient intervention strategies. However, the data are predominantly stored in free text in multiple large documents rather than in structured format, making unassisted human review and analysis unfeasible, thus underscoring the need for automated knowledge discovery techniques. Automated extraction of structured meaning from narratives to find, interpret, and prioritize knowledge, with a focus on identifying social history information such as substance use and mental illness, can provide a reprieve from the time- and cost-prohibitive nature of human review, allowing probation officers and personnel to devote their time to higher priority tasks that require human cognitive skills.

This article describes the application of a Natural Language Processing (NLP) and Artificial Intelligence (AI) system to discover important information on substance use, mental illness diagnoses, and treatment history of individuals under supervision. We developed a system for the automatic detection of four main events in the social history of the individual under supervision within free-text probation documents: (i) any evidence of substance use (alcohol, prescription, and illegal drugs) also defined as "indefinite diagnoses" for SUD; (ii) official diagnoses for SUD; (iii) official diagnoses for COD; and (iv) history of contract or non-contract treatment for SUD. We also automatically identify and extract related information (e.g., temporal information, facilities, treatment type, treatment outcome) within the text. The results are combined with metadata information from the Probation and Pretrial Services Automated Case Tracking System (PACTS) on client demographics and supervision dates. The system applies analytics to large data sets (N=98,389) over multiple documents (254,585 total documents and over 14 million Chrono entries), fuses the extracted information in a structured form, and performs analytic reasoning to enhance results. The results show that about 93 percent of this population have had a substance use issue in their lives and about 15 percent have officially been diagnosed with SUD, while approximately 29 percent have also received a formal diagnosis for a co-occurring disorder. Top mental disorders that co-occur with addiction are depressive disorders, anxiety disorders, bipolar I disorder, ADHD, and posttraumatic stress disorder (PTSD). Nearly 58 percent of

[1] The Center for Substance Abuse Treatment defines co-occurring disorders as the presence of one or more mental disorders as well as one or more disorders relating to the use of alcohol and/or other drugs. A diagnosis of COD occurs when at least one disorder of each type can be established independently of the other and is not simply a cluster of symptoms resulting from the one disorder (SAMHSA, 2020, p. ix).

these individuals have undergone substance use treatment at some point in their lives, while 35 percent have been under treatment for SUD or intended to attend treatment while under supervision.

The advantage of automated knowledge discovery from probation documents is that the system can access information beyond what is listed in structured form in PACTS and can be re-applied to new data sets or scaled up to process larger data sets, at no additional cost. There are several projects that have applied AI technology to the identification of substance use or mental conditions in unstructured clinical text, especially within clinical notes fields in Electronic Health Records (EHR), but only a few studies have targeted comorbidity detection. These approaches can automatically extract a range of information on these conditions, allowing researchers to process larger data sets than a manual review would allow. To our knowledge, however, the current investigation is the first study to use NLP and AI methods on probation narrative text to automatically detect mental conditions, substance use, and comorbidity issues among individuals on supervision in the criminal justice system.[2] The system also generates more detailed information than in previous clinical works.

Background

Cohort Studies and Surveys

Each year, the National Survey on Drug Use and Health (NSDUH) collects data on a wide range of behavioral health issues from a representative sample of U.S. adults, including those under court-ordered supervision within the past year, specifically on probation or parole. Based on these self-reported responses by 201,400 individuals, the 2012 study estimated that among males 18-49, 40.3 percent of probationers and 38.3 percent of parolees had an alcohol or illicit drug use disorder in the previous year. With respect to substance use treatment, nearly half of male probationers and parolees needed treatment; however, only about a quarter of probationers and less than one third of parolees received some treatment in the previous year. About 10 percent of probationers reported that they were receiving treatment at the time of the survey, and about 3 to 7 percent had received treatment in prison or jail within the year (SAMHSA, 2014). For male individuals over 50 on probation or parole, Bryson et al. (2019) find that 21 percent of participants in the NSDUH survey reported a serious or moderate mental illness within the past year, and about 80 percent reported receiving some sort of mental health treatment.

These statistics show that the number of probationers and parolees with mental or substance use disorders whose treatment needs are not being met by community treatment and supportive services is significant. Yet, statistics for individuals under federal supervision who present with both SUD and COD are not readily available. Based on information in PACTS, Mangione (2019) finds that for post-conviction supervision, federal probation offices supervised 186,509 cases during fiscal year 2018. Of that number, 120,217 (64 percent) had substance abuse treatment conditions. During the same period, federal probation offices had 27,122 persons (14.5 percent) in substance use contract treatment.[3] The study adds that individuals with co-occurring disorders receive substance use and mental health services in an integrated fashion, but it does not provide statistics on these individuals.

A few studies have focused on the relationship between SUD with co-occurring mental illness and rates of recidivism on a smaller scale. Magee et al. (2021) conduct a retrospective cohort study of all individuals arrested in 2016 in Indianapolis, Indiana (N=22,939), by linking their arrest information with their clinical mental health and SUD diagnoses in the two years before the arrest. They found that 27.7 percent of the individuals in the study were formally diagnosed with SUD and 22.5 percent also had evidence of COD. The authors also found that individuals with SUD or co-occurring conditions in the preceding 2 years are at higher risk of repeat arrest, and they advocate interventions aimed at low-level offenders with behavioral health needs to prevent recidivism. Constantine et al. (2012) reach a similar conclusion after using a retrospective cohort design to study rearrest rate of inmates with serious mental illness diagnoses[4] in the Pinellas County, Florida, jail between July 1, 2003, and June 30, 2004, and their health and social service data from 2002 to 2006 (N=37,236). They find that 10.1 percent of the inmates from that period met the criteria. The authors argue that individuals with serious mental illness, especially with co-occurring SUD diagnosis, are at higher risk of felony rearrest, compared with other populations of inmates.

As this section demonstrates, previous studies of individuals with SUDs and CODs in the justice system have predominantly been limited to small-scale analyses in cohorts in specific counties, performed mostly manually, and relying almost exclusively on administrative data, from which results are then extrapolated to the general population. Analysis at the federal level typically depends on self-reported responses to the NSDUH survey questions, which does not emphasize cooccurrences of SUD and COD among the probationer and parolee population. Mangione (2019) is the most directly relevant study, yet it is limited to structured data provided in PACTS and does not have access to information on non-contact treatments or the social history information captured in free text form in probation documents.

Automated Approaches

Several efforts have focused on using automated NLP techniques, including machine learning, to extract smoking or substance use status (e.g., "Past smoker," "Current smoker"). In these approaches, the authors typically use a pre-defined list of substance-related key phrases to identify text (paragraphs, sentences) containing potential substance use mentions from the notes, prior to processing the text through the NLP system. For instance, Uzuner (2008) describes several systems for classifying the smoking status of patients by using machine learning and rule-based algorithms, and reports F-scores ranging from 84 to 90.[5] Ni et al. (2021) developed an automated substance use detection system to identify substance use information in pediatric settings (N=3,890). Besides status (lifetime or current user), the system also detects substance

[2] Earlier versions of this study are described in Rowland et al (2018) and Megerdoomian et al (2019).

[3] "Contract" treatment refers to cases where Judiciary funds are used to pay facilities for treatment of individuals under supervision. PPSO will also frequently use treatment services that are available to the person under supervision in the community without cost to the federal Judiciary or through the individual's own healthcare coverage. This is referred to as "noncontract" treatment. (Mangione, 2019).

[4] Serious mental illness includes schizophrenia, schizoaffective disorder, bipolar I disorder, major depressive disorder, other psychotic disorders, and other bipolar and mood disorders.

[5] F-measure or balanced F-score is a measure that combines precision and recall (harmonic mean). Precision (also called positive predictive value) is the fraction of relevant instances among the retrieved instances, while recall (also known as sensitivity) is the fraction of relevant instances that were retrieved.

categories (tobacco, marijuana, alcohol, opiates, any use) and agent (if family member was the user or the participant). The authors compare a knowledge-based model using logic rules and regular expressions and a deep-learning model trained on pre-annotated data and find that the deep-learning model performs better on most substance use categories and assertions, with sensitivity of 87.5 percent and specificity of 89 percent,[6] while their knowledge-based model outperforms the deep-learning model in detecting opiates use.

A couple of studies focused on identifying patients at increased risk of problem opioid use by applying NLP to electronic health records for patients receiving chronic opioid therapy (COT). Hylan et al. (2015) study chronic noncancer patients starting COT (N=2,752). Their algorithm addresses linguistic variation (different words with the same meaning), polysemy (single words with several meanings), negation (e.g., "reports no pain" vs. "reports severe pain"), ambiguity, and temporality. The algorithm resulted in a sensitivity of 60.1 percent and specificity of 71.6 percent. Carrell et al. (2015) also applied NLP to identify clinician entered descriptions of problem opioid use in the unstructured clinical notes of patients. They also capture terms that are negated, as well as terms qualified by uncertainty, historical reference, or reference to a person other than the patient. The false positive rate for patients identified by the NLP system was 41 percent. Authors conclude that human-assisted review of results is important for validation purposes.

Few studies have targeted both mental illness and substance use detection. Ridgeway et al. (2021) applied NLP to unstructured text sections of clinical notes in EHRs for HIV patients to detect mental illness and substance use among people living with HIV (N=778). The system performs keyword search using a list of pre-defined indicative words, negation terms, and regular expressions developed by subject matter experts. The study identified high rates of mental illness and substance use among patients in an urban HIV care clinic, nearly half of whom did not have a diagnosis code in the structured patient records, suggesting that relying on structured EHR fields alone to identify people with behavioral health disorders may miss a substantial number of patients. The NLP algorithm for detecting mental illness had a Positive Predictive Value (PPV)(Precision) of 98 percent and a Negative Predictive Value (NPV) of 98 percent. The NLP algorithm for detecting substance use had a PPV of 92 percent and an NPV of 98 percent.[7]

Wang et al. (2015) are one of the few studies that apply more in-depth analysis beyond regular expressions and logic rules and extract a wider range of attributes related to substance use. The system detects three main sub-categories of substance use (alcohol, drug, and nicotine use), but also extracts more fine-grained elements including amount, frequency, type (e.g., wine, alcohol, tobacco), status (current, past), method, and temporal information. The authors developed a knowledge-based NLP system that leverages substance use lexicons and annotated linguistic resources along with deep dependency parse relationships between tokens provided by the Stanford Dependency parser (Manning et al., 2014). The authors report the F-scores of 89.8, 84.6, and 89.4, respectively for alcohol, drug, and nicotine use statement detection. Performance on the extraction of attributes report average F-scores of 82.1 (amount), 90.3 (frequency), 80.8 (status), 88.7 (method), 96.6 (type), and 74.5 (temporal). The lower score on the temporal attribute is due to the variability in expression with temporal expressions.

As this review of the literature shows, NLP technologies have been used for extracting a range of information on substance use and mental illness from clinical notes. These approaches detect entities that correspond to Unified Medical Language System (UMLS) concepts (e.g., drugs, diseases, medications, or procedures). More recent approaches also detect features like negation and the experiencer or subject. Temporal analysis is typically limited to basic expressions that identify the status of use (current, past, lifetime), although more recent approaches have integrated analysis of more complex temporal expressions. The current study goes beyond previous works by building a system that can identify both substance use and mental conditions and related concepts, as well as mentions of treatments and diagnoses, including non-contract cases that are only discussed in narrative text. Furthermore, the system automatically detects all important dates and determines the status of treatments by leveraging textual information. The application of NLP allows the processing of larger data sets, combining structured PACTS information with knowledge discovered from relevant textual documents, delivering a comprehensive study of substance use and comorbidity detection among individuals in the criminal justice system.

Methods

System Overview

The Advanced Narrative Analytics System Infrastructure (ANAnSI) performs content extraction and detailed narrative analytics for knowledge discovery within a distributed high-performance system infrastructure. ANAnSI is a hybrid system that leverages linguistic resources including substance use and mental condition lexicons, and combines them with probabilistic algorithms as well as knowledge-based analytics to identify and extract rich event-based narrative analysis at the sentence level (i.e., *who did what to whom, where, and when* analysis). The system also uses linguistic knowledge in machine algorithms to perform reasoning tasks (e.g., temporal reasoning) and integrates machine learning-based components to make data-driven predictions (e.g., treatment outcome analysis). ANAnSI processes each sentence in the data collection and produces a detailed event-based analysis. Additional domain-specific analysis discovers properties relevant to substance use and mental illness. Table 1 shows the types of features automatically detected for each event, while Table 2 illustrates sample SUD treatment sentences with their corresponding analytic representation.

System analysis can also be viewed in terms of the relationships between each event and its participants, including the subject or agent of the action, the person affected by the action or the patient, and relations to temporal expressions. Figure 1 illustrates the analysis for the sentence in Table 2 where the Begun_by and Ended_by relations capture the start and end dates of the event. Shaded boxes represent discovered elements and arrows represent relationships between these elements.

[6] Sensitivity (True Positive Rate) refers to the proportion of those who received a positive result on this test out of those who actually have the condition, and Specificity (True Negative Rate) refers to the proportion of those who received a negative result on this test out of those who do not actually have the condition (when judged by the "Gold Standard").

[7] PPV and NPV in the clinical domain allow one to say how likely it is for a patient to have a specific disease. The positive predictive value is the probability that following a positive test result, that individual will truly have that specific disease. The negative predictive value is the probability that following a negative test result, that individual will truly not have that specific disease.

Data

The study focuses on 98,389 probation individuals under supervision as of October 2021. The information for this project was automatically obtained from the free text sections of Presentence Investigation Reports (PSIR), which represent investigations into the history of the person convicted of a crime before sentencing to determine if there are extenuating circumstances. More recent information was collected from psychological assessments and reports from treatment providers, and Chrono entries where probation officers record notes on office or home visits with the individual under supervision. All extracted events are associated with the offender's jurisdiction, criminal offense, and demographic information in the database for easy search and retrieval. A breakdown of the corpora used is shown in Table 3.

To prepare the data for processing, preprocessing steps are required.

1. Select corpus files: The PSIR documents and Chrono entries are explicitly tagged as such in PACTS and are easy to identify. However, psychological assessments may be classified under different document types. We therefore created a set of heuristic rules to automatically determine which documents should be treated as psychological evaluations by performing a keyword search on the Notes section of the PACTS metadata where users indicate additional information about the document type. For instance, if the Notes section contained terms such as "psy", "eval", "stable", "abel", "eval", "evl", "treatment", "trt", "ass" (for assessment) and did not include the terms "contract", "waiver", "receipt", or "no show", the documents were selected for analysis.
2. Extract text from PDF documents: The system applies generalized content extraction to the scanned and electronic PDF documents associated with the individuals under supervision. In addition, this component performs document structure analysis on the Presentence Investigation Reports to identify and parse out the different sections of the PDF documents and extracts the tabular profile and criminal information as well as all free text content per section. The following sections of the PSIR are predominantly used to extract relevant domain information: Mental and Emotional Health, Substance Abuse, Personal and Family Data, Juvenile Adjudications, Employment History, Education and Vocational History, Adult Criminal Convictions, Criminal History, and the PSIR cover page. This component further "cleans" the data by normalizing the textual content to maximize processing.

TABLE 1
Automatically Extracted Substance Use and Comorbidity Related Indicators

Feature	Value / Examples
Event type	Formal diagnosis, Substance use, Treatment
Status	Negated, non-negated
Mental condition	Substance Use Disorder, Anxiety, Schizoaffective disorder, etc.
Substance	Ethanol, Cocaine, Opioids, Promethazine, etc.
Usage method	Snort, drink, smoke, etc.
Temporal	Date, range ("since 2000"), duration ("for two months"), frequency ("twice a week")
Reporter	Self-report, medical professional, (medical) records, third party
Spatial	Facility, Organization, Location
Treatment type	Inpatient, outpatient
Procedure	Counseling, mental health treatment, rehabilitation, substance abuse treatment, etc.
Treatment outcome	Completed, discontinued, ongoing, participated, terminated, intended

TABLE 2
Representation of Automated Analysis for Sample Sentences

Sample Sentence	*Structured Representation of Analysis*
"The defendant reported he was unsuccessfully discharged from intensive outpatient substance abuse program at The Pyramid Rehabilitation Center, St. Louis, Missouri on December 10, 2012."	***EventType:*** Treatment ***Status:*** Non-negated ***MentalCondition:*** Substance Use Disorder ***Facility:*** The Pyramid Rehabilitation Center ***Location:*** St. Louis, Missouri ***Date:*** 12-10-2012 ***Reporter:*** self ***TreatmentType:*** Outpatient ***Procedure:*** Substance abuse program ***TreatmentOutcome:*** Terminated
"Information indicates that he was seen by Dr. Joseph Smith from January 30, 2011, until August 12, 2013."	***EventType:*** Treatment ***Status:*** Non-negated ***MentalCondition:*** N/A ***MedicalPerson:*** Dr. Joseph Smith ***StartDate:*** 1-30-2011 ***EndDate:*** 8-12-2013 ***Reporter:*** third party ***Procedure:*** Treatment

FIGURE 1
Event, Entity, and Relations Analysis

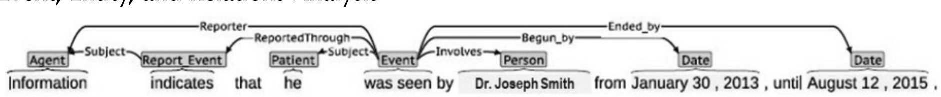


TABLE 3
Breakdown of Corpus for the Study

Documents Total	**9,721,569**
PSIR	121,518
Psychological Evaluation	133,067
Chronos	14,407,656
Sentences Total	**80,489,549**

N = *98,389*

Technical Approach

The ANAnSI architecture is provided in Figure 2. The *Information Extraction Component* takes the text extracted from PDF documents and Chrono notes as input and leverages open-source NLP tools for in-depth linguistic analysis and parsing. Stanford CoreNLP is a probabilistic system that performs entity recognition and sentence segmentation, detects the part-of-speech categories of each term, generates the dependency parse structure for each sentence, and detects temporal expressions (Manning et al., 2014). We used CoreNLP version 3.92 to generate dependency structures for all narrative statements. Apache cTAKES (clinical Text Analysis and Knowledge Extraction System) was developed specifically to extract and analyze clinical information from unstructured text (Savona et al., 2010).

The *Knowledge Discovery and Integration Component* builds on the results of the previous components to structure a complete analysis for events and their participants, to refine the recognition and classification of entities, and to infer the temporal relations between events in generating a timeline. This section also applies advanced linguistic analysis to improve argument and negation detection and improve precision of results. The complete list of entities employed by the system are person, role (e.g., doctor), location (city, state-or-province, country), facility, organization, temporal, money, salary. This component is able to detect and label the events in each sentence as well as related events such as a reporting event (*Medical records indicate that …*), an aspectual event that marks the end or beginning of the main event (*He began treatment in 1993*), or an event marking an intention (*Johnson stated that he would like to attend treatment for depression*). Finally, the system applies temporal reasoning techniques to infer complex temporal relations between events (e.g., *He became depressed after his infant brother died in 2000* infers that "became depressed" event began in or after the year 2000), compute temporal expressions with respect to the referred date (e.g., the diagnosis date can be computed based on the individual's date of birth in *He was diagnosed at age 20*, or the treatment date can be computed based on the document date in *She is currently undergoing treatment for anxiety* or *Medical records indicate that the defendant was terminated from the program three weeks ago*). In addition, this component links all temporal relations to obtain a complete temporal graph to capture the start and end dates of an event (e.g., *Jackson began treatment in March 2003 and was successfully discharged three months later* will be analyzed as having a start date of 2003-03-01 and an end date of 2003-06-01).

The resulting analyses are enhanced for the use case in the *Domain Analysis* component by identifying significant events and relations for the mental health and substance use domains. This component can detect paraphrases of mental condition mentions and substances at the appropriate level of granularity for PPSO. For example, terms such as *depression, chronic depression, depressive tendencies,* and *major depressive disorder* are all mapped to the more general term *depressive disorder.* In addition, it detects relevant events such as diagnoses, prescriptions, drug use, or treatments by identifying verbs commonly associated with these events (e.g., *attend, complete, hospitalize, undergo* are typically used to describe treatment events). This component also classifies the start or end of an event. For instance, verbs such as *enroll* or *enter* signal the beginning of a treatment event while *discharge* indicates the end of the treatment program. For each treatment event, the system also detects the treatment provider or facility mentioned in the sentence, the nature of treatment (e.g., inpatient or outpatient), and the procedure (e.g., anger management, drug rehabilitation). Any negated events are tagged as such.

As the system performs linguistic argument analysis to identify the participants in each event, it is possible to distinguish and ignore cases where a family member is mentioned rather than the individual under supervision (e.g., "*the defendant's mother suffered from Schizophrenia*"). The system also tags the source of the information (e.g., reported by a medical professional or self-reported). Note that a sentence like *The defendant denied smoking marijuana* is analyzed as a negated substance use event, but the reporter is tagged as "self-report." As descriptions of treatment events in probation documents can use very divergent wording to report the outcome, the system applies a trained Maximum Entropy machine learning algorithm to automatically classify the treatment outcome based on the categories shown in Table 4. The algorithm performs at 85 percent accuracy.

Extracted information from all documents associated with a particular individual is stored in the *Knowledge Model*, a Neo4j graph-based database management system, allowing all analyses to be compiled in a structured form with explicit links between related concepts and properties. The graph-based representation facilitates viewing all relevant information associated with a given individual, as well as obtaining an overview of all individuals under supervision in a specific district. The database provides the user with

FIGURE 2
Analytic Pipeline for ANAnSI

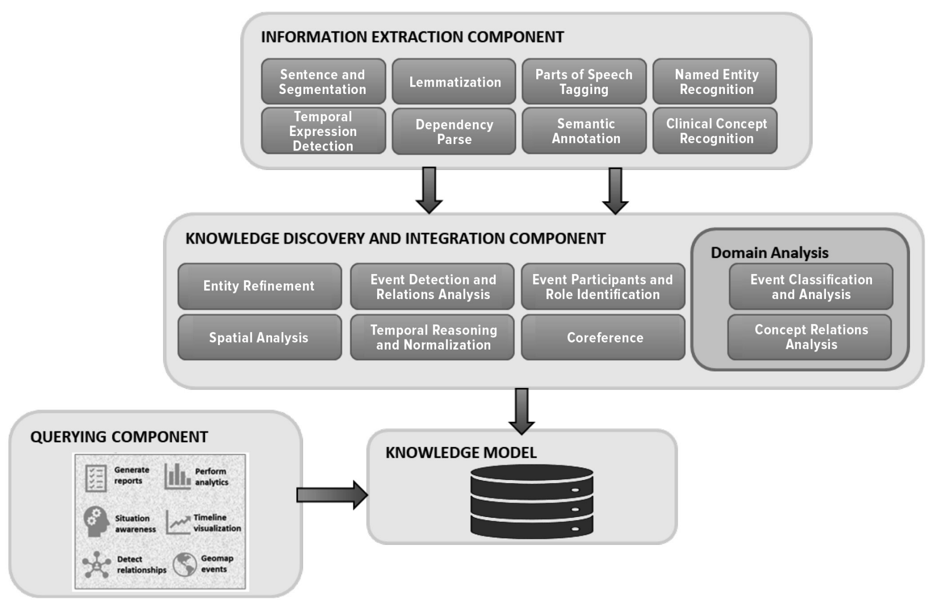


TABLE 4
Treatment Outcome Categories

Outcome Category	Description	Example
Participated	treatment occurred, but outcome not specified	"He received some substance abuse treatment while incarcerated as a teenager and as an adult."
Completed	treatment occurred and was successful	"He made 'good progress' in treatment and was successfully discharged on January 25, 2005."
Terminated	treatment occurred, but client was removed from the program before completion	"The defendant admitted he was discharged due to continued use while in treatment."
Discontinued	treatment occurred, but client left before completion	"He reported he was ordered into residential treatment but did not attend and absconded from supervision."
Ongoing	treatment is still taking place	"As a condition of his pretrial release, the defendant attends weekly drug treatment counseling sessions."
Intended	the client plans to enroll in treatment	"When asked, the defendant advised he would like to participate in the Residential Drug Abuse Program while incarcerated in the Bureau of Prisons."

a powerful query language to easily display answers to research questions, which can in turn be displayed in various formats, such as comma-separated value files, timeline view of events, or event relationship visualizations.

Results

Overview of Results

ANAnSI automatically processed over 62 million sentences from the documents associated with 98,389 individuals under supervision and identified events indicating formal SUD diagnoses, reports of substance use, presence of co-occurring disorders, and treatment participation information for the study group. Since recent events are of more interest to PPSO, the system also verifies if the detected event date or the document date falls within the supervision period for the individual. These results are shown in Figure 3 and details are presented in Table 5, contrasting all events in each class with the subset that was identified as occurring within the supervision period. We treat "successful" treatments as the ones that were not discontinued, terminated, or extended.

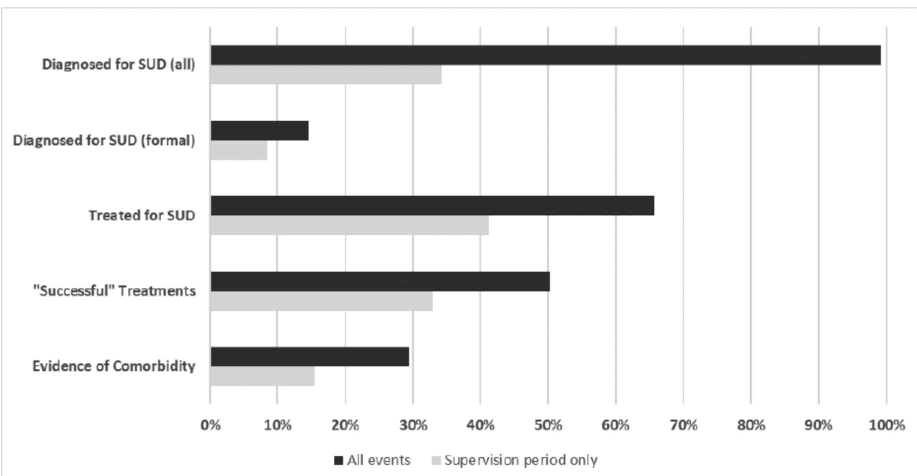
Results of the automated analysis show that about 93 percent of the individuals under supervision have had an issue with substance use at some point in their lives (see Diagnosed for SUD (formal & indefinite)). These include formal diagnoses for SUD as well as informal reports of substance use, either self-reported or reported by a medical professional or a third party. In contrast, only about 15 percent of the study group has received an official diagnosis at some point in their lives, according to the files analyzed. Note that these numbers reduce to 28 percent for all substance use reports and about 8 percent for formal diagnoses if limited to events within the supervision period. Furthermore, about 29 percent of the population under study has had at least one diagnosis for a co-occurring disorder with their substance use at some point in their life, and about 15 percent have received a COD diagnosis during their supervision period. Table 6 shows a comparison of the system results with the findings in previous research. As previously discussed, none of the previous study groups or research approaches correspond directly to the current analysis, yet a comparison of estimated values shows certain correlations.

The comparatively higher treatment scores found by ANAnSI reflect the treatment records during supervision as documented in the Chrono entries and include both contract and non-contract treatments. If we only take into account older information provided through PSIR or treatment reports, the percentage of individuals treated for SUD is lowered to 44 percent overall and 21 percent during their supervision period, which is closer to the rate discovered in previous studies. Similarly, the information on the individual's substance use history reflects data from the Chrono notes, while the information limited to PSIR and treatment documents finds that a smaller number (about 79 percent of the population under study overall and about 16 percent during the supervision period) has had a history with substance use.

Figure 4 illustrates the co-occurring disorders most often mentioned in the documents. These results correlate with findings in the literature where "mental disorders likely to co-occur with addiction include depressive disorders, bipolar I disorder, posttraumatic stress disorder (PTSD), personality disorders (PDs), anxiety disorders, schizophrenia and other psychotic disorders, ADHD, and eating and feeding disorders" (SAMHSA, 2020).

Meanwhile, Figure 5 shows the substances most often reported – about 74 percent of the individuals under supervision reported using alcohol (Ethanol), while 71 percent reported using marijuana (Marihuana) in their social histories or psychological evaluations. Other top substances include Cocaine, Methamphetamine, Ecstasy, and Heroin.[8]

FIGURE 3
Event Types Automatically Identified by ANAnSI



Percent values refer to the percentage out of the total number of clients (N=98,389).

TABLE 5
Results of Automated Event Detection for Substance Use and Co-occurring Mental Disorders

N=98,389 clients	All life events		During supervision period only	
	count	% of total	count	% of total
Diagnosed for SUD (formal)	14,321	14.56%	8,219	8.35%
Diagnosed for SUD (formal & indefinite)	91,077	92.57%	27,847	28.30%
Diagnosed for COD	28,288	28.75%	14,645	14.88%
Treated for SUD	57,507	58.45%	34,096	34.65%
Successful treatments	43,228	43.94%	26,581	27.02%

TABLE 6
Comparison of automated ANAnSI output with previous research findings

	SUD Diagnosis	Substance Use History	Treated for SUD	COD
ANAnSI (all)	14.56%	92.57%	58.45%	28.75%
ANAnSI (during supervision)	8.35%	28.30%	34.65%	14.88%
Magee et al., 2021	27.70%			22.50%
NSDUH, 2012	40.30%		25.00%	21.00%
Mangione, 2019		64.45%	14.54% (contract only)	
Sabol and Couture, 2008		60-80%		

[8] "Pharmaceutical Preparations" refer to general mention of the term "drugs" in the text.

FIGURE 4
Most Common Co-occurring Disorders Detected

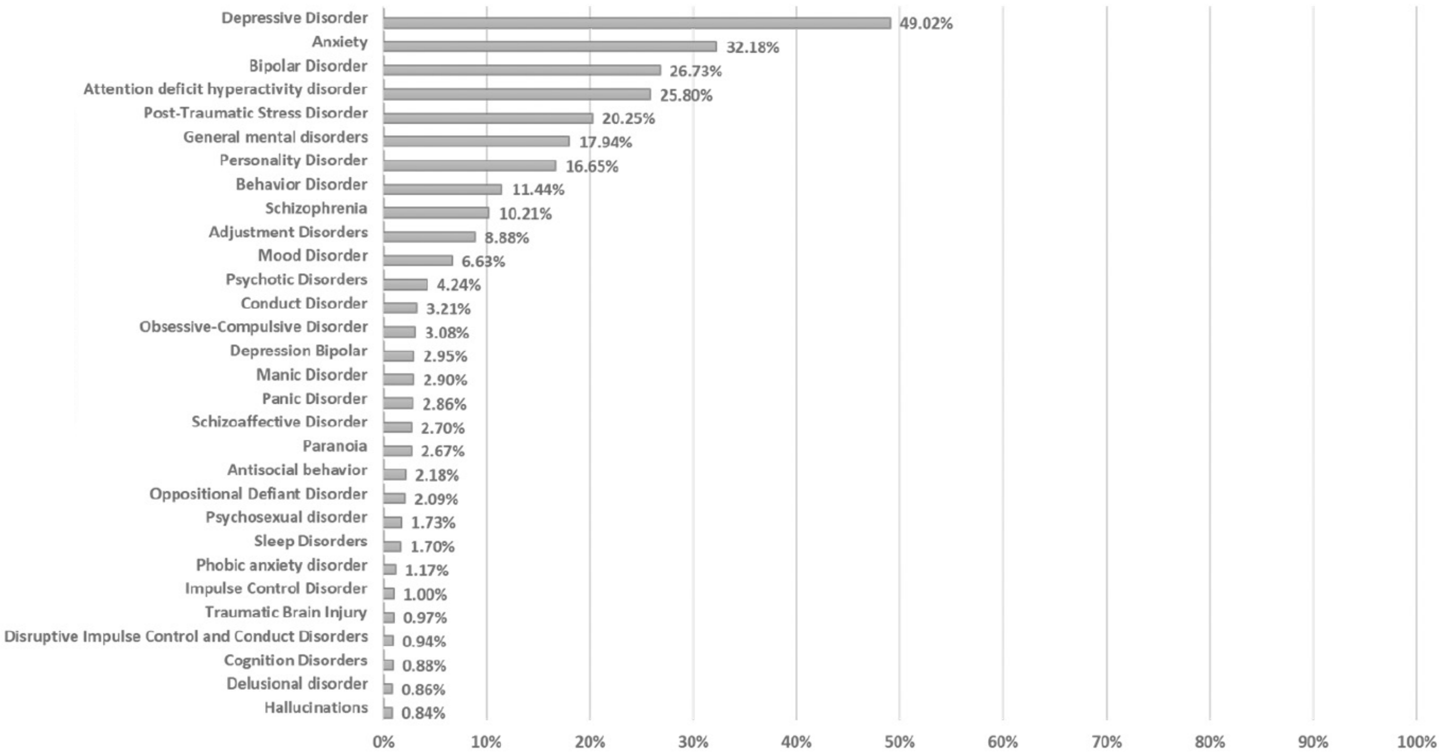
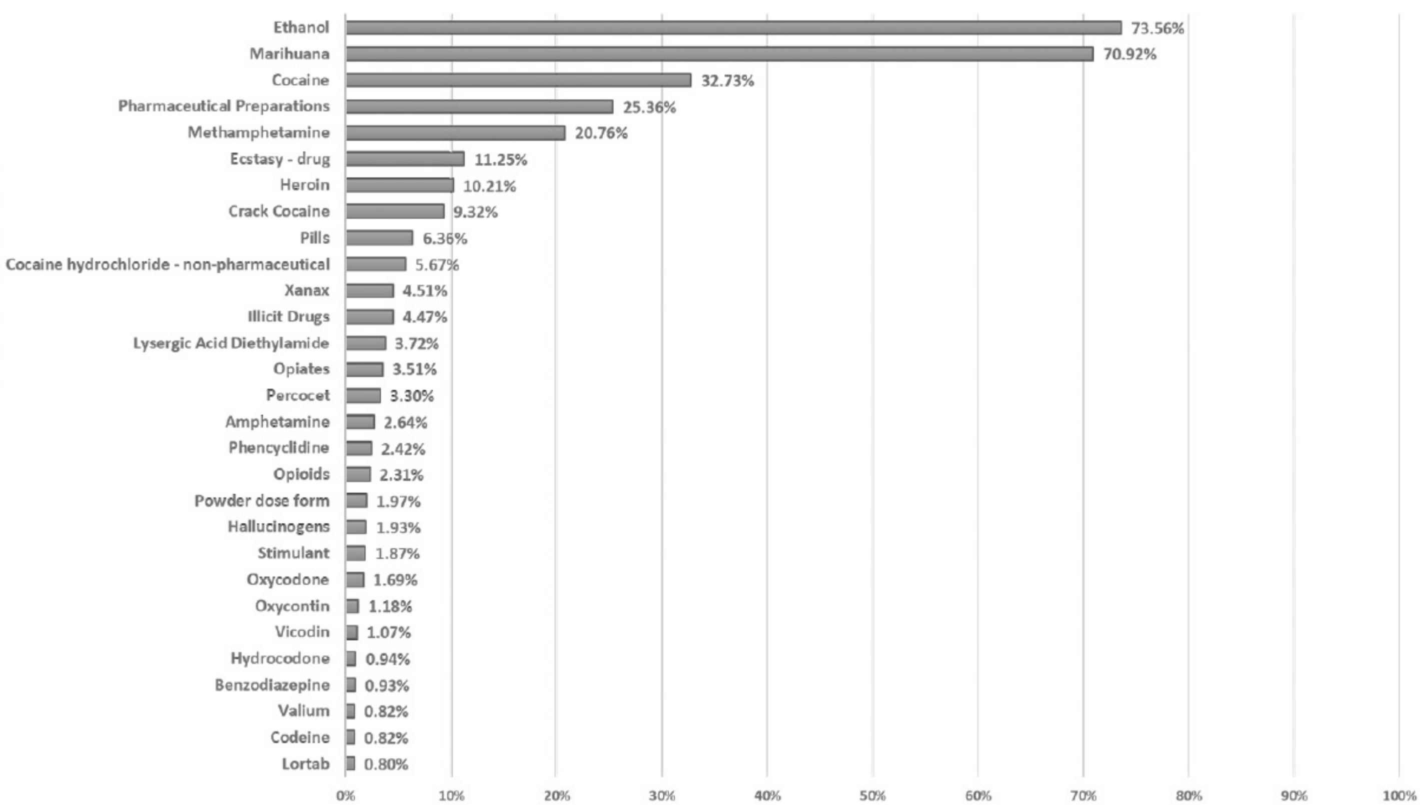


FIGURE 5
Percentage of Clients with SUD (all) that Used a Given Substance at Some Point in Their Life



A study of the top treatment procedures results in an expected list where the majority of the population has participated in substance abuse treatment, substance abuse counseling, or rehabilitation programs, with a smaller group participating in a mental health treatment.

Performance Results

System performance was evaluated by creating a small reference sample of about 500 sentences from PSIR and treatment reports to measure the accuracy of the information extracted for each event type. The 500 sentences were manually annotated by team members for all event types and event attributes of interest. The language analytics results were then compared to the pre-annotated reference set to measure how many of the detected elements were accurate and to also calculate how many of the expected elements were not picked up by the system. System performance does not fare as highly when applied to the analysis of Chronos, however, given the more informal writing style and content, which often lacks full sentences and contains various shorthand as well as misspellings. (See Table 7.)

While this study makes an important contribution to advancing methods to extract substance use and mental condition information from text, there are limitations that could be addressed in future enhancements. The system does not consider same or co-referring events or substances repeated in distinct sentences. For instance, the sentences "*The defendant began substance use treatment in August 2010. He was successfully discharged in May 2012*" will be analyzed as depicting two distinct treatment events, instead of merging them as a single treatment with a begin date of 2010 and end date of 2012. Another challenge that was left unaddressed in the current version of ANAnSI is the distinction between events (e.g., diagnoses, treatments) that occurred in the past and those that are currently valid. This can be accomplished by leveraging the tense and aspect information that the system computes to enhance detection accuracy.

TABLE 7
Performance Results

Events & Attributes	Recall	Precision	F-measure
Diagnose	100	92.86	96.30
Usage	93.94	100	96.88
Treatment	90.91	100	95.24
Mental condition	94.12	100	96.97
Reporter	75	84.38	79.41
Date	76	73.1	74.52
Subject of event	100	98.28	99.13
Treatment outcome	100	70	82.35
Treatment procedure	81.82	100	90.00

Conclusion

This article describes a successful approach to the automatic extraction and analysis of probation narrative text in the mental health and substance use domain. The results provide evidence that the use of technology in identifying important information in free narrative text in administrative records is feasible and cost-effective, and any adaptations to new domains can be accelerated through probabilistic methods. These analytics can be further developed in various directions, depending on the mission needs of the organization.

Acknowledgments

We would like to acknowledge the guidance and support of the U.S. Probation and Pretrial Services Office throughout this project. The project was led and funded by the U.S. Probation and Pretrial Services Office and the Technology Solutions Office at the Administrative Office of the U.S. Courts as part of the Applied Technology Research and Development program and managed by the Judiciary & Legal Technology Modernization Department at the MITRE Corporation.

This (software/technical data) was produced for the U.S. Government under Contract Number USCA16D2019, and is subject to Judiciary Policy Clause 6-60, Rights in Data—General (JUN 2012).

No other use other than that granted to the U.S. Government, or to those acting on behalf of the U.S. Government under that Clause is authorized without the express written permission of The MITRE Corporation.

For further information, please contact The MITRE Corporation, Contracts Management Office, 7515 Colshire Drive, McLean, VA 22102-7539 (703) 983-6000.

Data Availability Statement

The data underlying this article are from the U.S. Probation and Pretrial Services Office–funded project to study substance use and co-occurring disorders among individuals under supervision. The data cannot be shared publicly due to protected health information of individuals that participated in the study.

References

Bryson, W. C., Cotton, B. P., Barry, L.C., Bruce, M. L., Piel, J., Thielke, S. M., and Williams, B. A. (2019). Mental health treatment among older adults with mental illness on parole or probation. *Health and Justice* 7(4).

Carrell, D. S., Cronkite, D., Palmer, R., Saunders, K., Gross, D. E., Masters, E. T., Hylan, T. R., and Von Korff, M. (2015). Using natural language processing to identify problem usage of prescription opioids. *International Journal of Medical Informatics*, 84(12):1057-1064.

Constantine, R. J., Robst, J., Andel, R., and Teague, G. (2012). The impact of mental health services on arrests of offenders with a serious mental illness. *Law and Human Behavior*, 36(3): 170-176.

Drake, R. E., and Wallach, M. A. (1989). Substance abuse among the chronic mentally ill. *Hospital and Community Psychiatry*, 40(10):1041-6.

Fazel, S., Hayes, A. J., Bartellas, K. Clerici, M., and Trestman, R. (2016). Mental health of prisoners: Prevalence, adverse outcomes, and interventions. *Lancet Psychiatry*, 3(9):871-81.

Fearn, N. E., Vaugh, M. G., Nelson, E. J., Salas-Wright, C. P., DeLisi, M,. and Qian, Z. (2016). Trends and correlates of substance use disorders among probationers and parolees in the United States, 2002–2014. *Drug and Alcohol Dependence*, 167:128-139.

Feucht, T. E., and Gfroer, J. (2011). Mental and substance use disorders among adult men on probation or parole: Some success against a persistent challenge.

Hylan, T. R., Von Korff, M., Saunders, K., Masters, El, Palmer, R. E., Carrell, D., Cronkite, D., Mardekian, J., and Gross, D. (2015). Automated prediction of risk for problem opioid use in a primary care setting. *The Journal of Pain*, 16(4):380-387.

Magee, L. A., Ranney, M. L., Fortenberry, J. D., Rosenman, M., Gharbi, S., and Wiehe, S. E. (2021). Identifying nonfatal firearm assault incidents through linking police data and clinical records: Cohort study in Indianapolis, Indiana, 2007–2016. *Preventive Medicine,* 149.

Mangione, C. (2019). Overview of substance use disorder occurrence and treatment in the Federal Judiciary. *Federal Probaation*

Journal, 83(2).

Manning, C. D., Surdeanu, M., Bauer, J., Finkel, J., Bethard, S. J., and McClosky, D. (2014). The Stanford CoreNLP Natural Language Processing Toolkit. In *Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, pp. 55-60.

Megerdoomian, K., Branting, K., Horowitz, C. E., Marsh, A. B., Modly, N., Petersen, S. J., Scott, E. O., and Wariyar, S. B. (2019). Automated narrative extraction from administrative records. *Proceedings of the Workshop on Artificial Intelligence and the Administrative State (AIAS 2019)*, Montreal, QC, Canada.

Monahan, J. (1992). Mental disorder and violent behavior: Perceptions and evidence. *American Psychologist*, 47 (4), 511–521.

Ni, Y., Bachtel, A., Nause, K., and Beal, S. (2021). Automated detection of substance use information from electronic health records for a pediatric population. *Journal of the American Medical Informatics Association*, 28(10):2116-2127.

Ridgeway, J. P., Uvin, A., Schmitt, J., Oliwa, T., Almirol, E., Devlin, S., and Schneider, J. (2021). Natural language processing of clinical notes to identify mental illness and substance use among people living with HIV: Retrospective cohort study. *JMIR Medical Informatics,* 9(3).

Rowland, M. G. (2018). Federal probation and pretrial services: What's going on and where are we going? (Presentation by the Chief of PPSO). Probation and Pretrial Services Office, Administrative Office of the U.S. Courts.

Sabol, W., and Couture, H. (2008). Prison Inmates at Midyear 2007. *Prison and Jail Inmates at Midyear.* Bureau of Justice Statistics Bulletin, June 2008, NCJ 221944.

SAMHSA (2014). The NSDUH report: Trends in substance use disorders among males aged 18 to 49 on probation or parole. March 6, 2014. Retrieved at https://www.samhsa.gov/data/sites/default/files/sr084-males-probation-parole/sr084-males-probation-parole/sr084-males-probation-parole.htm.

SAMHSA (2020). Substance use disorder treatment for people with co-occurring disorders, TIP42.

SAMHSA (2022). Key substance use and mental health indicators in the United States: Results from the 2022 National Survey on Drug Use and Health. Retrieved at https://www.samhsa.gov/data/sites/default/files/reports/rpt42731/2022-nsduh-nnr.pdf.

Savova, G. K., Masanz, J. J., Ogren, P. V., Zheng, J., Sohn, S., Kipper-Schuler, K. C., and Chute, C. G. (2010). Mayo clinical Text Analysis and Knowledge Extraction System (cTAKES): Architecture, component evaluation and applications. *Journal of American Medical Informatics Association* 17: 507–513.

Stahler, G. J., Mennis, J., Belenko, S., Welsh, W. N., Hiller, M. L., and Zajac, G. (2013). Predicting recidivism for released state prison offenders: Examining the influence of individual and neighbourhood characteristics and spatial contagion on the likelihood of reincarceration. *Criminal Justice and Behaviour,* 40(6):690-711.

Uzuner, Ö, Goldstein, I., Luo, Y., and Kohane, I. (2008). Identifying patient smoking status from medical discharge records. *Journal of the American Medical Informatics Association*, 15(1):14–24.

Wang, Y., Chen, E. S., Pakhomov, S., Arsoniadis, E., Carter, E. W., Lindemann, E, Sarkar, I. N., and Melton, G. B. (2015). Automated extraction of substance use information from clinical texts. *AMIA Annual Symposium Proceedings*, 2121-2130.

Development and Testing of a Digital Coach Extender Platform for MOUD Uptake

Jessica Vechinski
Center for Health Enhancement & System Studies (CHESS), University of Wisconsin
Dharmaraj Veeramani
University of Wisconsin – Madison
Barbara Bowers
University of Wisconsin – Madison
Todd Molfenter
Center for Health Enhancement and System Studies (CHESS), University of Wisconsin

BOTH THE CRIMINAL legal system (CLS) and the health systems are complex, and their interagency relationships can further complicate effective dissemination, adoption, implementation, and sustainment of evidence-based practices and treatments, including the implementation of medications for opioid use disorder (MOUD). Coaching is a favored implementation strategy,[1,2] but it is labor intensive for the coach, the organization, and the involved staff. This is a substantial barrier and often makes this pivotal implementation strategy costly, particularly in human resources. Accordingly, coaching techniques need to be designed for scaling up and affordability to maximize the full potential of the external coaching function. Researchers at the University of Wisconsin–Madison and George Mason University under the Justice Community Opioid Innovation Network (JCOIN) funded by the National Institutes of Health (NIH)'s (U2CDA050097, MPI Taxman and Rudes), Helping End Addiction Long Term (HEAL) Initiative are conducting a pilot that will include development of a Coaching Extender Platform (CEP). CEP is an asynchronous communication approach that does not require live or synchronous communication between the coach and the site. CEP's objective is to provide an affordable way to extend the coaching function and increase coaching effectiveness. The pilot has two study aims: 1) Design and develop the CEP prototype using user-based needs assessment and user-centered design strategies and Web application development best practices and 2) Conduct a six-month pilot with four jail settings to assess CEP's ability to increase targeted MOUD use and to understand the factors that promote or impede CEP implementation.

Introduction

The Centers for Disease Control (CDC) identifies overdose prevention as a national priority and expanding access to addiction treatment services as essential to responding to the opioid overdose epidemic.[3] Nearly 11 million individuals pass through local jails yearly,[4,5] 5 million people are on parole or probation,[6] and 1.5 million people are in state and federal prisons.[4] The Criminal Legal System (CLS) has a constitutionally driven responsibility to provide behavioral health care (i.e., mental health and substance use services) to this large concentration of U.S. adults with behavioral health needs; however, less than 10 percent of justice-involved individuals are able to access behavioral health services regardless of setting (jail, probation, etc.).[7-10] Among CLS populations, 66 percent have substance use disorder (SUD),[3] 15 percent identify with lifetime opioid use,[3] and 11 percent are pain medication dependent.[3] These symptoms and use rates are dramatically elevated compared to those of the general population, resulting in unfavorable rates of overdoses,[11] suicide,[12-14] disabilities and physical disorders,[11,15,16] homelessness,[17] and death.[18,19]

The three most common medications for opioid use disorder (MOUD)—methadone, injectable naltrexone, and buprenorphine—have all proven to increase retention in treatment and decrease self-reported use of opioids, criminal activity, and mortality.[8] While pharmacotherapy holds excellent promise, medications are underused in SUD treatment, both in and out of the CLS.[20,21] Approximately 80 percent of those with opioid use disorder (OUD) do not receive appropriate treatment.[9,10] Use of MOUD among CLS populations is even lower,[22] with justice-referred patients being one-tenth as likely to receive agonist MOUD as other patients.[23] This inequity is particularly unwarranted as individuals in incarceration settings have direct access to health care, sometimes for the first time in their lives, and are ten times more likely to die because of an overdose post-incarceration.[24,25] Individuals in the state of New York receiving buprenorphine or methadone treatment for OUD during incarceration were associated with an 80 percent reduction in overdose mortality risk for the first-month post-release.[26] However, despite the promise of MOUD, their rates of use have remained persistently low in CLS settings.[27,28]

Strategies are needed for increasing low MOUD rates in jail settings that can address the complexities of CLS health systems and resistant CLS personnel attitudes towards MOUD.[27] External coaching from someone independent of the organization has become a standard strategy for behavior and systems change[29-31] and has resulted in significant improvements in evidence-based practice implementation,[32,33] administrative functions,[31,34] clinical processes,[35,36] and systems of care.[37,38] Coaching is identified as an active ingredient in learning collaboratives and has been one of the more successful implementation strategies.[34] However, a significant deficit with coaching is its labor intensiveness for the coach, the organization, and the involved staff. In times of labor crisis, this becomes a substantial barrier and often makes this pivotal implementation strategy cost- and human resources-prohibitive. Accordingly, coaching techniques need to be designed for scaling up and affordability to maximize the full potential of the external coaching function. Moreover, greater clarity and consistency regarding what occurs within the coaching sessions is needed. This "black box" of coaching results in variation in practice and, consequently, in results overall. In a current trial conducted by this research team in jail settings through a Justice Community Opioid Innovation Network (JCOIN) initiative,[39] the promise and limitations of coaching became prominent, motivating the team to attempt to develop a coaching approach that could optimize the benefits of coaching while overcoming the strategy's limitations.

The Parent JCOIN Study

Researchers at the University of Wisconsin–Madison and George Mason University under the JCOIN initiative through the National Institutes of Health (NIH), Helping End Addiction Long Term (HEAL), are conducting an implementation effectiveness trial with 42 jails and community-based treatment provider organizations around the nation that are working to adopt, implement, or increase buprenorphine, methadone, and injectable naltrexone MOUD programming within their correctional setting.[5] The study "Fostering MOUD Use in Justice Populations" is in year four of a five-year study that began in January of 2021. The study is looking at two different implementation strategies, NIATx Coaching and ECHO, to determine the optimal approach for increasing the uptake of MOUD. NIATx coaches provide technical assistance in MOUD implementation and organizational change to assist justice and treatment organizations in implementing and disseminating MOUD for justice clients. ECHO focuses on the MOUD provider's knowledge and self-efficacy of MOUD care to increase confidence in using MOUD[40] through monthly telementoring sessions. Sites were randomly assigned to one of four study arms that compared low-dose NIATx Coaching (4 one-hour coach calls in one year) and high-dose NIATx Coaching (12 one-hour coach calls in one year) with and without ECHO. The study hypothesizes that sites assigned to the study arm, including high-dose NIATx coaching and ECHO, will be most successful in implementing or expanding MOUD use. The focus on implementation was on enhancing the MOUD Cascade of Care (CoC) of *screening*, *identification*, *referral*, *medication administration*, and *community transition*.

During the study, change team members were invited to participate in one-hour semi-structured qualitative interviews at the end of the intervention phase of the study to learn how coaching and ECHO (if applicable) impacted their site's MOUD programming at both the organizational and personal level and their experiences with receiving coaching and ECHO. These interviews provided great insight into the barriers and benefits of providing coaching strategies within a complex environment such as the criminal legal setting. One recurring theme that presented itself was an overwhelming request to communicate more with their assigned coach between coach calls and have a more asynchronous or timely communication method to ask questions, receive feedback, and keep each other informed of the process improvements happening within the site. The feedback led the research team to devise ways to bridge this communication gap and, ultimately, the beginning steps of designing a coaching platform that is structured, asynchronous, and digital to provide an affordable way to extend the coaching function and increase coaching effectiveness without increasing labor intensiveness.

The use of online technology through a laptop or tablet to expand access to and improve the coaching function will be developed and tested through the Coaching Extender Platform (CEP). The CEP will rely on online asynchronous communication and will initially be used with a limited amount of live coaching. The CEP and live coaching "hybrid model" will be designed as an implementation strategy that facilitates the application of other Expert Recommended Strategies for Implementing Change (ERIC),[41] such as conducting education sessions, identifying and preparing champions, developing and organizing quality monitoring systems, conducting cyclical tests of change, and audit and feedback.

Development and Assessment of the Coaching Extender Platform

The development and assessment of the CEP pilot has two aims: 1) Design and develop the CEP prototype using user-based needs assessment and user-centered design strategies and Web application development best practices, and 2) Conduct a six-month pilot with four jail settings to assess CEP's ability to increase targeted MOUD use and understand the factors that promote or impede CEP implementation.

Aim 1: Development of a Coaching Extender Platform Prototype

User-centered design (UCD) is a fundamental approach in software development that places the end users at the heart of the design and development process.[42] It is a methodology that prioritizes the needs, preferences, and feedback of users to create software that functions effectively and delivers superior user experience.[42] UCD recognizes that a successful software platform is one that aligns with the goals and requirements of its target audience. This approach involves a series of iterative stages that encompass understanding, designing, and evaluating the user's interactions with the software. By continuously involving users throughout the development lifecycle, UCD seeks to ensure that the resulting software is intuitive, efficient, and capable of meeting the users' specific needs. User personas, user journey maps, and wireframing are key practices within the UCD approach that play a pivotal role in crafting user-friendly software. Software products developed with these tools tend to result in higher user satisfaction and adoption rates[43] (Figure 1).

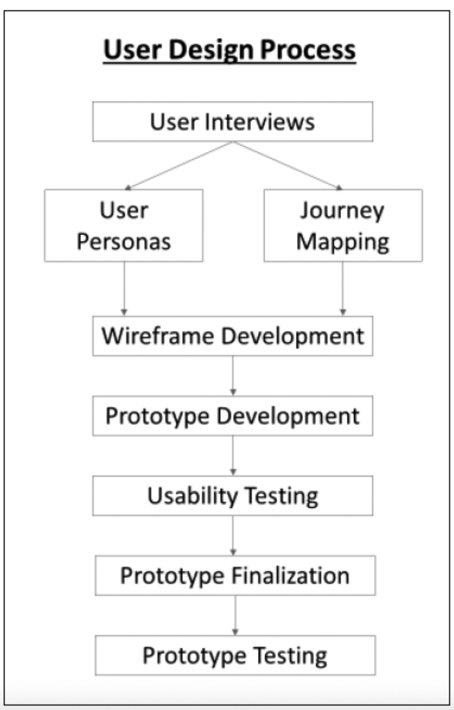
Creating User Personas

User personas are detailed profiles representing various segments of the target user base that provide a valuable reference point during the design and development phases.[44,45] The user perspective helps teams empathize with users and make informed decisions related to software functionality and interface. Creating user personas is a methodical process that involves conducting comprehensive user research (e.g., through qualitative interviews) to gather data on demographics, behaviors, needs, and preferences. From this research, commonalities and patterns among users are identified and detailed persona profiles are crafted, giving each persona a representative identity, including information such as goals, challenges, and technological proficiency. These personas are then prioritized based on their relevance to the software's goals and shared with the development team to foster empathy for the target users and guide user-centered design decisions throughout the development process.

Journey Mapping

User journey mapping lays out the entire user experience, from the initial interaction with the software to the completion of user-intended tasks. This visualization aids in identifying and prioritizing user needs and expectations and helps anticipate potential points of friction in the user experience and opportunities for improvement.

FIGURE 1
User-Centered Design Process



Wireframing

Wireframing involves creating skeletal outlines of the user interface, illustrating the layout and functionality of the software. It serves as a blueprint, facilitating early testing and validation of design concepts by the software development team. The wireframe is used to design the prototype.

Interviews with Target Users

The persona development and journey mapping methodology began with interviews designed to gain clinical and workflow-relevant insights from individuals working in the jails and providing MOUD coaching to the jails. This UCD approach collected data through a structured interview with four NIATx coaches with jail MOUD implementation experience and eight jail staff. Participants were asked questions about their beliefs, successes, challenges, and practices in promoting MOUD in jail settings and their past experiences using NIATx change methods. Example questions for jail staff included: What are your primary job responsibilities? Can you describe your typical day and the activities you perform? What are your main frustrations and pain points in your role of providing MOUD? How have you benefited from the live coaching sessions? What would be the goals of using a coaching platform? What are the top three functionalities you would look for in a coaching platform? Example questions for coaches included: What need(s) do you see filling for those you are coaching? What are your needs that you feel are unmet or underserved by live coaching sessions? Why and how might you use a web-based coaching platform? What would be your expectations and anticipated benefits of using a platform? What are the top 3 functionalities you would look for in a coaching platform?

Data from each participant was entered into a spreadsheet and aggregated separately for coaches and jail staff. A qualitative descriptive approach was used to aggregate and describe the participants' responses and to review the variation and commonality of responses.[46] Although the responses showed some differentiation between the groups, the two groups had consistent overall themes.

Interview Findings

From the interviews, three types of personas emerged and were the starting point for creating the CEP wireframe: 1) the *NIATx Coach*, 2) the *MOUD Executor* (nurse practitioner, physician, program manager, and/or lieutenant/sergeant/sheriff responsible for the day-to-day MOUD program), and 3) the *Executive Champion* (medical or program director overseeing the MOUD programming and funding).

The research team compiled the qualitative findings into a matrix for each persona, including 1) barriers experienced in providing MOUD, 2) goals for using CEP to overcome barriers, and 3) desired outcomes. All three personas identified seven barriers to providing MOUD in a correctional setting, with the most prominent barrier being the lack of communication between staff at each level of the jail. (The CLS setting is a complex system with processes structured around standard operational procedures and guidance from multiple stakeholders—often with little direct correlation or communication between the two.) A close second was stigma associated with MOUD, not only from leadership and staff, but also from those incarcerated. A jail may have strong leadership support of MOUD, but if staff carrying out the program's day-to-day operations are not in support, the program fails. Similarly, if leadership is not in favor, regardless of staff receptiveness, the program will not succeed. Other barriers identified were limited staff bandwidth, inadequate funding to provide MOUD, lack of community treatment provider partnerships, and inefficient tracking and monitoring of the MOUD cascade of care (number screened, referred, administered medication, and referred to treatment post-release).

Following the identification of barriers, interviewees were asked what they would find helpful in the CEP that would assist in combating the barriers. For all three personas (NIATx Coach, MOUD Executor, & Executive Champion), responses were unanimous that having more asynchronous communication between the coach and site would be beneficial. The recurring message from the MOUD Executor and Executive Champion Personas was that the NIATx coach kept their site on track with process improvement projects, was a motivator and a sounding board for ideas, and validated their goals and missions. However, they felt they could have been more successful if communication with the coach had occurred more than monthly or quarterly during a scheduled coach call. The NIATx coaches echoed the same sentiment: if they had continuous updates on the jail's process improvement project(s) and were able to answer lingering questions or offer suggestions in a timely manner, they, too, would be able to provide more effective coaching. All three personas also relayed that it would be useful to have one organized, central location to house agendas, task lists, process improvement charter forms, and MOUD data so that at any given time, either the coach or jail staff could get a quick status update on progress, pull reports from the MOUD data for yearly reports or funding applications, and/or review information that was discussed in prior communications, rather than searching through an email inbox.

The MOUD Executor and Executive Champion Personas presented a few other prominent themes. Interviewees suggested that it would be beneficial to interact with other jail staff/medical teams to discuss pertinent MOUD topics such as screening processes, medication administration protocols, and tactics for addressing stigma and diversion. Additionally, they indicated that it would be extremely helpful to learn about barriers other jails face and how they address those obstacles. Another theme was the need for more educational resources on MOUD (protocols, posters, papers, training/informational videos, and podcasts relating to MOUD) for staff and those incarcerated. Many jails are now mandated to provide one or more forms of MOUD but are not given adequate resources to easily implement or expand their programming. Providing resources that have worked with other jails can be a simple yet effective way to bridge the informational and skill development resource gap.

The NIATx Coaches shared one additional suggestion that was not raised by the others: an organizational needs assessment to be completed by the jail at baseline and throughout the coaching relationship to monitor progress. Coaches relayed that having a deeper understanding of the jail's organizational structure, the approach to implementing MOUD, and the barriers the site was encountering before coaching began would have allowed for more effective and efficient use of time in guiding the sites through their process improvement projects.

The information compiled through the interviews and User Persona development was then integrated into a user journey map that identified the software features to inform Wireframe development.

Creation of Wireframe and Prototype

Using the information compiled from the interviews, the team used a web-based design software to develop a wireframe (mock-up) of the CEP platform that was shared with developers. The wireframe included a mix of digital features that CHESS (Center for Health Enhancement & System Studies) platforms developed at the University of Wisconsin–Madison and that have previously been found beneficial in behavior change[47] along with new features and stylistic preferences generated from the qualitative interviews. It was determined that the alpha version of the CEP prototype would include the following features.

Project Management Center: Jail staff and coaches will use this feature to generate and store agendas and project charter forms, complete organizational needs assessments, manage tasks, and track progress towards implementation objectives. Both coaches and sites can view and comment on the information entered by the platform users. Automatic notifications will be built into this feature to notify the intended recipient(s) that information has been updated.

Cascade of Care (CoC) Performance Tracker: This feature will allow the jail to enter their CoC data (number screened, number referred, number of individuals administered medication, number of medication slots administered—buprenorphine, methadone, and naltrexone—and number of those referred to a community treatment provider post-release) and view it in an easy-to-understand graphical format that will also compare their CoC performance to that of similar jails.

Communication Center: Two message boards will be available. The first board will be for the coach and site to communicate progress made in enhancing MOUD programming and elements of the CoC. Sites can ask coaches for advice at any time. The second board will be for sites to communicate with one another to pose questions or share resources. Automatic notifications will be built into both discussion boards to notify the intended recipient(s) that a message or response has been posted.

Resource Center: Resources will be made available to enhance each of the steps in the CoC and will include an instant library (information), such as peer-reviewed articles that support different CoC practices, personal stories (how others have made improvements to the CoC), common policies and operating procedures, podcasts and informational videos, handouts, and funding opportunities.

Skills Toolbox: This feature will provide tutorials on applying different organizational change tools, including improving system linkages between jail and community care settings. Role-specific tools will be available for the executive sponsor, change leader/site liaison, and project team members.

Usability Testing

This phase will focus on getting feedback on CEP's alpha version of the prototype before it is tested in the pilot. The platform will be shared with the study team, NIATx coaches, and jail staff participating in the qualitative interviews. The project coordinator will conduct usability walkthroughs with jail, coach, and study team users. They will be able to navigate around the platform and view the features of CEP and perform a set of common tasks. Subsequently, they will be asked to provide feedback on the functionality, ease of use, and perceived usefulness. The information gathered will then be shared with the development team to refine CEP before the pilot's launch.

Aim 2: Conduct a Six-Month Pilot

A six-month pilot with four jails will study CEP's effectiveness in expanding coaching access and impact and achieving improved and more consistent implementation results. Four jail sites interested in expanding MOUD CoC programming will be recruited for the pilot. The pilot will be a "hybrid" coach design, including the asynchronous CEP and low-dose live synchronous coaching.

Pilot Activities

The CEP pilot will follow a project-focused design that begins with a one-hour online orientation focused on how to use and benefit the most from the different CEP features. The sites will also participate in a two-hour kick-off meeting where they will meet their assigned NIATx coach and research team, receive training on the NIATx change model, review expectations and requirements of the pilot, and begin working with their coach to identify their site's first process improvement project focused on MOUD programming while using CEP. The change leader and change team at each site will work with their assigned NIATx coach on one or more process improvement projects focused on implementing or improving their MOUD programming with the use of the CEP and participate in two one-hour coach calls at three months and six months. Throughout the six months, the CEP will be available to team members, and the change leader will be asked to interact with the CEP on a weekly basis.

Pilot Study Evaluation Plan

We will employ a pre-post evaluation plan, with data being collected at baseline and M6. The primary outcome measure will be MOUD use via methadone, buprenorphine, and injectable naltrexone. Secondary outcome measures will be cascade of MOUD care infrastructure (via Jail Substance Use Treatment Services Inventory),[48] Staff Attitudes toward MOUD,[49] NIATx Fidelity,[50] and Workplace Stress.[51] At the conclusion of the pilot, qualitative interviews will be conducted to allow jail and coach participants to describe their personal experiences using the CEP and whether or not CEP helped alleviate the barriers discussed in the initial qualitative interviews, including lack of communication, limited access to resources, and not having access to a network of other jails providing MOUD. Interviews will include specific, closed-ended questions to examine whether/how the CEP was used, whether/how CEP usage changed over time, how CEP contributed to achieving study outcomes, and how the CEP integrated with the live coaching function. The qualitative results will be used to assess platform effectiveness, including factors promoting and undermining the success of CEP during the pilot, and to enhance the CEP for future applications.

Conclusion

The public health imperative of providing MOUD in incarceration settings, where infrastructure complexities and stigma towards MOUD persist, provides a challenging and opportune setting to test the CEP. The CEP pilot will provide researchers and the development team with the necessary information to gain initial insights into the utility of virtual coach supports and evaluation feedback on how to refine CEP for effective use on a larger scale. The CEP's intended purpose is to promote scaling up and affordability of coaching to maximize the full potential of the external coaching function to address the opioid crisis and other pressing public health issues.

References

1. Molfenter T, Kim JS, Zehner M. Increasing engagement in post-withdrawal management services through a practice bundle and checklist. *The Journal of Behavioral Health Services & Research.* 2021;48(3):400-409.
2. Rutkowski BA, Gallon S, Rawson RA, et al. Improving client engagement and retention in treatment: The Los Angeles County experience. *Journal of Substance Abuse Treatment.* 2010;39(1):78-86.
3. Fearn NE, Vaughn MG, Nelson EJ, Salas-Wright CP, DeLisi M, Qian Z. Trends and correlates of substance use disorders among probationers and parolees in the United States 2002–2014. *Drug Alcohol Depend.* 2016;167:128-139. doi:10.1016/j.drugalcdep.2016.08.003
4. Kaeble D, Glaze L. Correctional Populations in the United States, 2015. 2016. https://www.bjs.gov/index.cfm?ty=pbdetail&iid=5870. Accessed February 12, 2019.
5. Minton T, Zeng Z. Jail Inmates in 2015. 2016. https://www.bjs.gov/content/pub/pdf/ji15.pdf. Accessed February 13, 2019.
6. Kaeble D, Bonczar T. Probation and Parole in the United States, 2016. https://www.bjs.gov/content/pub/pdf/ppus15.pdf. Accessed February 13, 2019.
7. Schwartz RP, Gryczynski J, O'Grady KE, et al. Opioid agonist treatments and heroin overdose deaths in Baltimore, Maryland, 1995-2009. *Am J Public Health.* 2013;103(5):917-922. doi:10.2105/AJPH.2012.301049
8. Mattick RP, Breen C, Kimber J, Davoli M. Buprenorphine maintenance versus placebo or methadone maintenance for opioid dependence. *Cochrane Database Syst Rev.* 2014;(2). doi:10.1002/14651858.CD002207.pub4
9. Saloner B, Karthikeyan S. Changes in Substance Abuse Treatment Use Among Individuals With Opioid Use Disorders in the United States, 2004-2013. *JAMA.* 2015;314(14):1515-1517. doi:10.1001/jama.2015.10345
10. Wu L-T, Zhu H, Swartz MS. Treatment utilization among persons with opioid use disorder in the United States. *Drug Alcohol Depend.* 2016;169:117-127. doi:10.1016/j.drugalcdep.2016.10.015
11. Bronson J, Maruschak L, Berzofsky M. Disabilities Among Prison And Jail Inmates, 2011-12. 2015. https://www.bjs.gov/index.cfm?ty=pbdetail&iid=5501. Accessed February 13, 2019.
12. Hayes LM, Rowan JR. *National Study of Jail Suicides: Seven Years Later.* Alexandria Va: National Center on Institutions and Alternatives; 1988.
13. Charles DR, Abram KM, Mcclelland GM, Teplin LA. Suicidal Ideation and Behavior Among Women in Jail. *J Contemp Crim Justice.* 2003;19(1):65-81. doi:10.1177/1043986202239742
14. Sarchiapone M, Jovanović N, Roy A, et al. Relations of psychological characteristics to suicide behaviour: Results from a large sample of male prisoners. *Personal Individ Differ.* 2009;47(4):250-255. doi:10.1016/j.paid.2009.03.008
15. Mallik-Kane K, Visher CA. Health and Prisoner Reentry: How Physical, Mental, and Substance Abuse Conditions Shape the Process of Reintegration. 2008:82. doi:10.1037/e719772011-001
16. Fazel S, Baillargeon J. The Health of Prisoners. *The Lancet.* 2011;377(9769):956-965. doi:10.1016/S0140-6736(10)61053-7
17. McNiel DE, Binder RL, Robinson JC. Incarceration Associated With Homelessness, Mental Disorder, and Co-occurring Substance Abuse. *Psychiatry Serv.* 2005;56(7):840-846. doi:10.1176/appi.ps.56.7.840
18. Binswanger IA, Stern MF, Deyo RA, et al. Release from Prison — A High Risk of Death for Former Inmates. *N Engl J Med.* 2007;356(2):157-165. doi:10.1056/NEJMsa064115
19. Merrall ELC, Kariminia A, Binswanger IA, et al. Meta-analysis of drug-related deaths soon after release from prison: Drug-related deaths after release from prison. *Addiction.* 2010;105(9):1545-1554. doi:10.1111/j.1360-0443.2010.02990.x
20. Stein BD, Sorbero M, Dick AW, Pacula RL, Burns RM, Gordon AJ. Physician Capacity to Treat Opioid Use Disorder With Buprenorphine-Assisted Treatment. *JAMA.* 2016;316(11):1211-1212. doi:10.1001/jama.2016.10542
21. Volkow ND, Frieden TR, Hyde PS, Cha SS. Medication-assisted therapies—tackling the opioid-overdose epidemic. *N Engl J Med.* 2014;370(22):2063-2066. doi:10.1056/NEJMp1402780
22. Brinkley-Rubinstein L, Zaller N, Martino S, et al. Criminal justice continuum for opioid users at risk of overdose. *Addict Behav.* 2018;86:104-110. doi:10.1016/j.addbeh.2018.02.024
23. Krawczyk N, Picher CE, Feder KA, Saloner B. Only One In Twenty Justice-Referred Adults In Specialty Treatment For Opioid Use Receive Methadone Or Buprenorphine. *Health Aff (Millwood).* 2017;36(12):2046-2053. doi:10.1377/hlthaff.2017.0890
24. Krawczyk N, Mojtabai R, Stuart EA, et al. Opioid agonist treatment and fatal overdose risk in a state-wide US population receiving opioid use disorder services. *Addiction.* 2020;115(9):1683-1694.
25. Binswanger IA, Stern MF, Deyo RA, et al. Release from prison--a high risk of death for former inmates. *N Engl J Med.* 2007;356(2):157-165.
26. Lim, S., Cherian, T., Katyal, M., Goldfeld, K. S., McDonald, R., Wiewel, E., ... & Lee, J. D. (2023). Association between jail-based methadone or buprenorphine treatment for opioid use disorder and overdose mortality after release from New York City jails 2011–17. *Addiction*, *118*(3), 459-467.
27. Scott, C.K., Grella, C.E., Dennis, M.L. *et*

al. Availability of best practices for opioid use disorder in jails and related training and resource needs: findings from a national interview study of jails in heavily impacted counties in the U.S.. *Health Justice* 10, 36 (2022). https://doi.org/10.1186/s40352-022-00197-3

28. Jail and prison opioid Project – medication for opioid use disorder in the criminal legal system (JPOP) 2022. http://prisonopioidproject.org/
29. Kitson AL, Rycroft-Malone J, Harvey G, McCormack B, Seers K, Titchen A. Evaluating the successful implementation of evidence into practice using the PARiHS framework: theoretical and practical challenges. *Implement Sci.* 2008;3:1.
30. Stetler CB, Legro MW, Rycroft-Malone J, et al. Role of "external facilitation" in implementation of research findings: a qualitative evaluation of facilitation experiences in the Veterans Health Administration. *Implement Sci.* 2006;1:23.
31. McCarty D, Gustafson DH, Wisdom JP, et al. The Network for the Improvement of Addiction Treatment (NIATx): enhancing access and retention. *Drug and Alcohol Dependence.* 2007;88(2):138-145.
32. Molfenter T, Kim H, Kim JS, et al. Enhancing Use of Medications for Opioid Use Disorder Through External Coaching. *Psychiatr Serv.* 2022:appips202100675.
33. Roosa M, Scripa JS, Zastowny TR, Ford Ii JH. Using a NIATx based local learning collaborative for performance improvement. *Evaluation and Program Planning.* 2011;34(4):390-398.
34. Gustafson DH, Quanbeck AR, Robinson JM, et al. Which elements of improvement collaboratives are most effective? A cluster-randomized trial. *Addiction.* 2013;108(6):1145-1157.
35. Hoffman KA, Ford JH, Choi D, Gustafson DH, McCarty D. Replication and sustainability of improved access and retention within the Network for the Improvement of Addiction Treatment. *Drug Alcohol Depend.* 2008;98(1-2):63-69.
36. Molfenter T. Reducing appointment no-shows: going from theory to practice. *Substance Use & Misuse.* 2013;48(9):743-749.
37. Molfenter T, Connor T, Ford J, Hyatt J, Zimmerman D. Reducing psychiatric inpatient readmissions using an organizational change model. *WMJ.* 2016;115(3):122-128.
38. Molfenter T, Kim JS, Zehner M. Increasing engagement in post-withdrawal management services through a practice bundle and checklist. *The Journal of Behavioral Health Services & Research.* 2021;48(3):400-409.
39. Molfenter T, Vechinski J, Taxman FS, Breno AJ, Shaw CC, Perez HA. Fostering MOUD use in justice populations: Assessing the comparative effectiveness of two favored implementation strategies to increase MOUD use. *J Subst Abuse Treat.* 2021;128:108370.
40. Komaromy M, Duhigg D, Metcalf A, et al. Project ECHO (Extension for Community Healthcare Outcomes): A new model for educating primary care providers about treatment of substance use disorders. *Substance abuse : official publication of the Association for Medical Education and Research in Substance Abuse.* 2016;37(1):20-24.
41. Powell BJ, Waltz TJ, Chinman MJ, et al. A refined compilation of implementation strategies: results from the Expert Recommendations for Implementing Change (ERIC) project. *Implement Sci.* 2015;10:21.
42. Interaction Design Foundation - IxDF. (2016, June 5). *What is User Centered Design?*. Interaction Design Foundation - IxDF. https://www.interaction-design.org/literature/topics/user-centered-design
43. C. K., Holden R. J., Valdez R. S. Human factors engineering and user-centered design for mobile health technology: Enhancing effectiveness, efficiency, and satisfaction. In: Duffy, V.G., Ziefle, M., Rau, PL.P., Tseng, M.M. (eds) Human-Automation Interaction. Automation, Collaboration, & E-Services, vol 12. Springer, Cham. https://doi.org/10.1007/978-3-031-10788-7_6
44. Billestrup J, Stage J, Bruun A, Nielsen L, Nielsen KS. Creating and using personas in software development: Experiences from practice. In: Sauer, S., Bogdan, C., Forbrig, P., Bernhaupt, R., Winckler, M. (eds) Human-Centered Software Engineering. HCSE 2014. Lecture Notes in Computer Science, vol 8742. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-44811-3_16
45. Wallach D, Scholz S. C. (2012). User-Centered Design: Why and How to Put Users First in Software Development. In: Maedche, A., Botzenhardt, A., Neer, L. (eds) Software for People. Management for Professionals. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-31371-4_2
46. Sandelowski M. Focus on Research Methods, Whatever Happened to Qualitative Description. *Research in Nursing & Health*. 2000;23:334-340.
47. Gustafson DH, Shaw BR, Isham A, Baker T, Boyle MG, Levy M. Explicating an evidence-based, theoretically informed, mobile technology-based system to improve outcomes for people in recovery for alcohol dependence. *Substance Use & Misuse.* 2011;46(1):96-111.
48. Scott CK, Dennis ML. (2019) Justice Community Opioid Innovation Network Jail Interview Protocol. Lighthouse Institute, Chestnut Health Systems. Manuscript Under Review.
49. Knudsen HK, Ducharme LJ, Roman PM, Link T. Buprenorphine diffusion: The attitudes of substance abuse treatment counselors. JSAT. 2015;29(2):95-106.
50. Gustafson D, et al. The NIATx Model: Process Improvement in Behavioral Health. Madison, WI: University of Wisconsin-Madison. 2011.
51. Gomez R, Summers M, Summers A, Wolf A, Summers J. depression Anxiety Stress Scales-21: Measurement and structural invariance across ratings of men and women. Assessment. 2014;21(4):418-426.

CONTRIBUTORS
To This Issue

Barbara Bowers
Professor, University of Wisconsin-Madison. Ph.D., University of California at San Francisco. Co-author of "Devaluing Incarcerated Populations: Deprioritizing Incarcerated Populations for COVID-19 Vaccination." *Western Journal of Nursing Research, 45*(11)(2024).

Ashit Chandra
Vice-President of Infobahn Solutions specializing in Data Engineering, Data Analytics, and Enterprise Data Management, designing and implementing Enterprise Data Lakes, Enterprise Data Warehouses, Business intelligence, and Big Data Analytics solutions that empower organizations to harness their data assets and meet mission goals.

Thomas H. Cohen
Senior Social Scientist, Probation and Pretrial Services Office, Administrative Office of the U.S. Courts. Previously, Statistician, Bureau of Labor Statistics. Ph.D., Rutgers University School of Criminal Justice. J.D., University of Maryland School of Law. Co-author of "Examining Federal Pretrial Release Trends Over the Last Decade," *Federal Probation* (Sept. 2018).

Hadeel Elyazori
Bachelor's degree, American University of Beirut. Ph.D. student and Graduate Teaching Assistant, Department of Information Science and Technology, George Mason University, Virginia.

Charles Horowitz
M.S.E.E., George Mason University. Previously, Principal Engineer, the MITRE Corp.

Kevin Lybarger
Assistant Professor, George Mason University. Ph.D., University of Washington. Co-author of "Leveraging natural language processing to augment structured social determinants of health data in the electronic health record." *Journal of the American Medical Informatics Association*, 2023. doi: 10.1093/jamia/ocad073. [free-access link] (April 2023).

Amy Marsh
Lead Computational Linguist, the MITRE Corp. M.S., University of Washington.

Karine Megerdoomian
Principal Artificial Intelligence Engineer, the MITRE Corp. Ph.D., University of Southern California. Previously, Research Scientist, University of Maryland, College Park. Co-author of "Induction of narrative models for legal case elicitation" (2023).

Todd Molfenter
Research Professor, Industrial and Systems Engineering, University of Wisconsin, Center for Health Enhancement and System Studies (CHESS). Ph.D., University of Wisconsin-Madison. Co-author of "Cascade of Care in the Legal System: Best Practices and Goals for Agencies Providing Care to Patients Concurrently Navigating Substance Use Disorders and Criminal Legal Involvement," *Journal of Drug Issues* (August 2023).

Matthew G. Rowland
Principal in Maloney, Rowland, & Associates LLC, providing diverse consulting services to government agencies, think tanks, and global technology firms focused on improving criminal justice outcomes. Previously, chief of the Probation and Pretrial Services Office for the Administrative Office of the United States Courts.

Amit Shah
President and Chief Data Officer of GNS-AI, working with various companies and government agencies in delivering safe, cost-effective, and innovative solutions. Previously, Data Science Manager at Abbott Laboratories. Ph.D. in Biomedical Engineering at the University of Illinois at Chicago. He led clinical studies on the application of virtual reality and adaptive feedback algorithms on neuro-rehabilitation of stroke and traumatic brain injury survivors.

Faye Taxman
Professor, George Mason University. Ph.D., Rutgers University-Newark. Author of *The Routledge Handbook on Global Community Corrections* (June 2024).

Teneshia Thurman
Researcher and post-doctoral fellow, George Mason University. Ph.D., George Mason University. Author of "Fidelity Measures for Risk-need Assessment Tools Usage in Case Plans" (2020).

Jessica Vechinski
Program Coordinator, University of Wisconsin, Center for Health Enhancement & System Studies (CHESS). M.S.W., University of Wisconsin-Milwaukee. Co-author of "Cascade of Care in the Legal System: Best Practices and Goals for Agencies Providing Care to Patients Concurrently Navigating Substance Use Disorders and Criminal Legal Involvement," *Journal of Drug Issues* (August 2023).

Dharmaraj Veeramani
Professor, University of Wisconsin-Madison. Ph.D., Purdue University. Co-author of "POMDP-based optimal maintenance planning with time-dependent observations," *European Journal of Operational Research*, Vol. 311, Issue 2, December 2023, 533-544, https://doi.org/10.1016/j.ejor.2023.05.022.

Federal Probation Style Sheet

A Journal of Correctional Philosophy and Practice

Federal Probation—published three times a year by the Probation and Pretrial Services Office, Administrative Office of the U.S. Courts—has been in print for more than 80 years. Articles published in the journal reflect current thought, practice, and research in the community supervision, corrections, and criminal justice fields. Authors include criminologists, academicians, researchers, and legal specialists, but also "front-line" probation, pretrial services, and corrections officers.

U.S. probation and pretrial services officers constitute the bulk of the approximately 5,500 recipients of *Federal Probation* in hard copy. Other readers include federal judges; correctional and criminal justice practitioners at the federal, state, and local levels; and educators and students. Most of our outside subscriptions are for university libraries; in addition, *Federal Probation* is available and widely accessed online at www.uscourts.gov.

The editors seek material that is either directly related to the work and professional interests of the readers or of significant peripheral interest. Articles of interest include those that describe or evaluate a new or innovative program or approach, discuss how a policy evolved, offer insight into an issue, propose a change, or provide historical perspective. The journal publishes articles on a wide spectrum of topics; articles in recent issues have focused on topics as diverse as risk assessment, evidence-based practices in community supervision, pretrial detention, disparity, prisoner reentry, location monitoring, officer safety training, and program implementation.

The editors of *Federal Probation* are looking for original material. Manuscripts submitted to *Federal Probation* should not be previously published or under simultaneous consideration elsewhere.

Authors preparing and submitting manuscripts should:

- In general, keep manuscript length at about 3500–5000 words, although we do publish longer articles, depending on the topic and treatment.
- Double-space manuscripts, allowing standard margins.
- Submit two or three title suggestions, and a brief abstract of your article.
- Type all quoted matter of more than three lines single-spaced and indented on both sides.
- Use tables and figures intentionally, and cite them in the text. That is, they should be used to support, illuminate, or substantiate the author's narrative argument or presentation.

Email to: Ellen Wilson Fielding, Editor, *Federal Probation* (email: Ellen_Fielding@ao.uscourts.gov), Administrative Office of the U.S. Courts, One Columbus Circle NE, Washington, DC 20544 (ph: 202-502-1651).

United States Government

INFORMATION

Order Processing Code

***5876**

Federal PROBATION

a journal of correctional philosophy and practice

Please send me _____ subscription(s) to *Federal Probation* at $16.50 each ($22.40 foreign) per year. The total cost of my order is $______________.

Price includes shipping and handling and is subject to change.

____

(PLEASE TYPE OR PRINT) NAME OR TITLE

____

COMPANY NAME ROOM, FLOOR, OR SUITE

____

STREET ADDRESS

____

CITY STATE ZIP CODE + 4

____

DAYTIME PHONE INCLUDING AREA CODE

____

PURCHASE ORDER NUMBER (OPTIONAL)

Method of Payment

❑ Check payable to: Superintendent of Documents

❑ GPO Deposit Account ______________________________

❑ Credit Card ❑ Visa ❑ MasterCard ❑ Discover

____

CARD NUMBER EXPIRATION DATE

____

AUTHORIZED SIGNATURE

ORDER ONLINE FROM

http://bookstore.gpo.gov

FAX YOUR ORDER TO

202-512-2104

PHONE YOUR ORDER TO

202-512-1800 or 866-512-1800

MAIL YOUR ORDER TO

Superintendent of Documents
P.O. Box 979050
St. Louis, MO 63197-9000

IMPORTANT!

Please include your completed order form with your remittance

thank you for your order!

FEDERAL PROBATION

Administrative Office
of the United States Courts

Washington, DC 20544