



Inside this box you'll find:

CAPACITY ISSUES

ACCESS INEQUITY

COSTLY PROCESSES

OBSCURE SYSTEMS

ANTIQUATED TECHNOLOGIES

ARDUOUS REQUEST PROCESSES

INCONSISTENT DATA RECORDING PRACTICES

A Look Inside the Black Box of New York State's Criminal Justice Data

Table of Contents

Executive Summary	1
Introduction	2
The Data Landscape: Key Themes	4
Data “Black Boxes”	5
Inconsistent, Antiquated Data Recording Practices	6
Arduous and Costly Processes	9
Access Inequity	12
Moving Forward: Where do we go from here?	16
Rethink the Purpose of Data Collection	17
Establish Data Standards	17
Address Ad Hoc Transparency	18
Alleviate Capacity Issues	19
Conclusion	20

Executive Summary

“The people’s right to know the process of governmental decision-making and to review the documents and statistics leading to determinations is basic to our society. Access to such information should not be thwarted by shrouding it with the cloak of secrecy or confidentiality.”

THE NEW YORK FREEDOM OF INFORMATION LAW¹

After experiencing a series of hurdles obtaining and analyzing criminal justice data in New York State, Measures for Justice (MFJ) set out to better understand the state’s data infrastructure. Drawing on interviews with system stakeholders--including practitioners, policy advocates, and researchers--we explored the quality and availability of criminal justice data in the state of New York. With heated criminal justice reform debates underway, there is a clear need for data that can speak to system performance. And yet our investigation uncovered that, with few exceptions, the mechanisms for criminal justice data collection and release in New York State are broken.

Efforts to put data to use across the state are frequently hampered by obscure systems, antiquated technologies, arduous request processes, and a degree of partiality that allows data access to some and not others. The present report explores each of these themes and ultimately suggests four pathways forward for New York agencies looking to pursue equitable and responsible data practices.

¹ See [NY Pub Off L § 84](#).

Introduction:

As debates about the effectiveness and fairness of New York’s criminal justice system wage on, the need for reliable and timely performance data is as apparent as ever.

Without access to information about the ways cases are processed through the system, practitioners, advocates, and researchers are left unequipped to make decisions necessary for their work. Moreover, the public is left in the dark about basic information pertaining to their local justice system.

Measures for Justice first set out to study our home state’s criminal justice system in 2014 and since then has submitted a series of data requests. After multiple attempts to acquire trial court data from the state, we could populate only a [limited number](#) of measures spanning the five boroughs of New York City. The question was why? Why were we unable to bring the same degree of transparency to New York as we did to, for instance, [Wisconsin](#), [Pennsylvania](#), [Virginia](#), and so many others? And if our experience is common as a criminal justice data and research organization, what does this mean for members of the general public, who have a right to understand what transpires in their state’s criminal justice system?

This report addresses the ways New York State’s criminal justice data infrastructure fails to meet basic levels of transparency that are requisite for evidence-based decision making and general accountability. It begins with an overview of New York’s main sources of centralized case-level criminal justice data, which constitutes the information associated with an individual’s criminal case. Next, we discuss four prominent themes that speak to the quality and accessibility of criminal justice data in New York State. We conclude by considering the impact of these findings and suggest four areas--data collection, standards, accessibility, and resourcing--to address in order to advance equitable criminal justice data practices throughout New York State. It is important to note that this report does not intend to serve as a comprehensive inventory of all data that may exist throughout the state. Instead, we hope this qualitative assessment promotes discussion about the importance of transparency for New York stakeholders, and provides a few starting points for criminal justice agencies looking to improve their data practices.

² See [MFJ’s Data Portal: New York](#).

What are New York State's main sources of criminal justice data?

Criminal case processing data are collected by a number of local and state entities throughout New York. At the local level, police, prosecution, and defense entities each collect data to some extent. These data collection and sharing practices vary extensively among, and often even within, local jurisdictions. At the state level, the most commonly referenced sources of centralized data include:

- The Division of Criminal Justice Services (DCJS);
- The Office of Court Administration (OCA); and
- The Department of Corrections Community Supervision (DOCCS)

DCJS

The Division of Criminal Justice Services (DCJS) serves as New York State's central repository for criminal history information. Whenever someone is arrested, with the exception of certain non-fingerprintable offenses,³ the data are transmitted from local law enforcement to DCJS. The DCJS database is linked with the case management systems of other state agencies so that as arrests move through the legal system, DCJS is equipped with the most recent case status. In addition, DCJS is responsible for the development of "eJusticeNY Integrated Justice Portal,"⁴ a portal designed to give qualified agencies access to current information for criminal cases throughout the state.

DOCCS

The New York State Department of Corrections and Community Supervision (DOCCS) collects information on all people subject to its jurisdiction. DOCCS is required to share certain data with DCJS and report annual statistics, such as facility capacity and staffing numbers, to the state legislature.

OCA

The Office of Court Administration (OCA), the administrative arm of the Unified Court System (UCS), oversees the state's Supreme, County, City, and District Courts, and maintains information related to cases originating in these places. Justice Courts ("town and village courts"), of which there are over 1,200, are the exception. Although these local courts are required to report to OCA select data elements from cases involving more serious offenses, data pertaining to non-fingerprintable offenses, such as traffic infractions and violations, are decentralized, living with each respective court.

³ Some offenses of lesser seriousness, such as traffic infractions and violations, are not required to be reported to the criminal history repository and are therefore considered "non-fingerprintable".

⁴ For more information, see <https://www.criminaljustice.ny.gov/ojis/ejusticeinfo.htm>.

The Data Landscape:

Key Themes

MFJ set out to gain a better understanding of the criminal justice data landscape in New York State from the perspective of the people who know it best. We began by reviewing state law and departmental policies governing access to criminal case information. Additionally, we drew information from documents, such as data sharing agreements, obtained from state agencies, and conducted a series of interviews with criminal justice stakeholders throughout New York State. We ultimately interviewed 36 people⁵ from a variety of backgrounds including researchers, policy advocates, and practitioners at both the state and local levels. It should be noted that several interviewees asked for anonymity because they feared retaliation. Interviews were roughly an hour long and semi-structured, beginning with a standard set of questions but left open for interviewees to lead the conversation in the direction they felt most pertinent based on their experience.⁶ What data exists? In what format? Is it accessible? Ultimately, four key themes emerged from this work.

4 Themes:

- I. Data “Black Boxes”**
- II. Inconsistent, Antiquated Data Recording Practices**
- III. Arduous and Costly Processes**
- IV. Access Inequity**

⁵ We spoke with 36 total people across 28 interviews. Some interviews included multiple respondents.

⁶ While MFJ sought to accurately reproduce the experiences described, respondents were not required to provide documentation of data collection efforts referenced.

Data “Black Boxes”

One of the themes that surfaced most frequently during our interviews was that stakeholders, including practitioners, policy advocates, and researchers, often do not know what information is being tracked and by whom. Several people we spoke with described data “black boxes” and futile attempts to learn about the information contained in case management systems across the state.

For public defender Tim Donaher, the mystery of which data are being tracked has prevented him from examining a number of questions of interest. Not only has Donaher struggled to obtain case-level data, his attempts to understand the types of data tracked by courts and other agencies have been unsuccessful. He once even resorted to requesting an agency’s internal software manual, hoping to better understand the system’s data fields and format, but the agency declined to share it. In another county, a local prosecutor expressed similar frustration when discussing their office’s attempts to access court data, “We’ve tried for several years to get access to the data but we can’t. We can’t even say, ‘Can you show us the available queries?’, ‘Can you just tell us what you’re collecting so we can tell you what we want?’ Honestly, we’re not even sure what OCA collects.”

“I’m absolutely certain that there are things they’re collecting that they may not share, or we don’t know to ask for.”

NYCLU Genesee Valley Chapter Director Iman Abid-Thompson described her experience requesting data at the local level: “I’m asking for information but have no idea what’s being collected or how long it’s being collected.” She discussed her efforts to obtain data via the state’s Freedom of Information Law (NY-FOIL). According to Abid-Thompson, oftentimes when her requests are denied agencies cite inadequate servers, but since the system is a mystery, there’s no way to confirm the existence of those inadequacies. A researcher working in the prison reform space described a similar experience in trying to obtain prison and parole information from the DOCCS data systems, noting, “I’m absolutely certain that there are things they’re collecting that they may not share, or we don’t know to ask for.” An advocate working on behalf of incarcerated

clients agreed, explaining the challenges this poses to holding officials accountable: “They just exist in this black box and act with impunity,” they explained.

In addition to the challenges this creates for data seekers, there is good reason to consider the difficulties this creates for agencies. Requests for information that agencies do not have creates a burden on the administrators charged with assessing and responding to these incoming requests. In walking us through the DCJS research request process, Deputy Commissioner Leigh Bates explained that part of the application review for case-level information is determining whether DCJS even has the data to answer the requestor’s questions, noting that requestors may not know what data DCJS has available to share for research purposes. Given that the agency may receive dozens or hundreds of requests for information each year, merely managing and responding to them equates to a lot of administrative work.

It’s easy to imagine how this lack of information might hinder efforts to access data. If you don’t know the data elements that agencies are tracking, you’re missing crucial information that can help you initiate a request and successfully acquire that data. This widespread opacity generates confusion for stakeholders working across the system, leaving people questioning which data are stored where and what can be reasonably accessed. If individuals who spend their careers working in the criminal justice system struggle to locate information about the information tracked, what could that mean for the general public’s ability to understand the data, much less obtain them themselves?

Inconsistent, Antiquated Data Recording Practices

Beyond the mystery around what data are collected, most of the stakeholders we interviewed noted concerns about the quality of the data they have been able to access. One of the most commonly cited problems relates to inconsistent data recording practices.

“ Different attorneys have different definitions, different people enter it, and no one checks it. ”

From advanced in-house case management systems like the one implemented at the Manhattan District Attorney's Office, to paper-based systems characteristic of smaller law enforcement agencies upstate, the ways data are defined and recorded vary significantly across jurisdictions and between agency types.

Respondents working in the prosecution realm noted how messy data look throughout the state. "Different regions speak differently. Actually, even the counties speak differently than other counties," one District Attorney explained. Even within the same office it can be difficult to ensure consistent and accurate data entry. As a former New York City prosecutor put it, "A prosecutor's office has to define its data, designate someone to enter the data, and then have someone audit the data. But oftentimes offices don't do any of those things. Different attorneys have different definitions, different people enter it, and no one checks it."

“Several interviewees reported counting file folders by hand.”

After inconsistency, antiquated data practices were the most commonly cited contributor to overall poor data quality. Several interviewees described counting file folders by hand and sifting through paper records to transcribe the information necessary for their analyses.

Other times, agencies may have digitized records, but rely on old technology that restricts the amount of, or ways in which, data are collected. Data Scientist Evan Misshula shared one example from his experience working with pretrial detainees at the New York City Criminal Justice Agency (CJA), an organization that aims to reduce unnecessary pretrial detention through evidence-based research and initiatives:

“We might look at the data and see that someone had their bail reduced to one dollar. In reality this was a flag to indicate [the person] had satisfied bond but there was a warrant in another place, but since there was nowhere to enter that in the database, people entering data relied on this ‘one dollar trick’. If you didn't know what that one dollar signified, you wouldn't know what's really going on...There's a tremendous amount of old data infrastructure and code that should be modernized.”

Again, it's important to note that these challenges don't affect all jurisdictions equally. While outdated practices hinder statewide analyses and county comparisons, it's important to note that New York City was often referenced as having advanced technological capabilities relative to the rest of the state. We revisit this in the report's final section.

Although the above limitations appear to impact the data generally, race data were noted as exceptionally challenging to work with. Practitioners and researchers alike explained how a reliance on officer perceptions of race often results in inaccurate data and ultimately impedes efforts to understand how people of different racial groups experience the system. Tina Luongo, Attorney-in-Charge at the Legal Aid Society in New York City gave one example. Drawing on their experience working with client data, they explained how people who identify as Hispanic are often underreported because an officer will assess their skin color as light and report them as white, or dark, in which case they're reported as Black.

One researcher who studies police misconduct described the discrepancies on individual rap sheets, explaining how the race specified at the top of a given criminal history record often contradicts what you find when looking at the breakdown by arrest. Additionally, race data recording methods vary across, and often even within, agencies, with some recording perceived race and others asking people to self-identify. Knowing whether race data are based on an officer's assessment or self-reported is vital for meaningful analysis of equitable justice, but the distinction is rarely made.

“They’re not necessarily tracking that data.”

Similar practices have shaped the available data on gender, making it difficult to assess the ways gender identity interacts with the criminal justice system. “Transgender folks in the system are classified as such but only to the extent they’re in a particular facility,” explained Felicia Henry, the Director of Research and Policy at the Correctional Association of New York (CANY). “So, trans women in Bedford Hills⁷ will be tracked there but we know trans folks are held at other institutions too, and they’re not necessarily tracking that data.” New York is certainly not unique in this way, and MFJ has found similar challenges in states across the country.⁸ Still, it’s important to highlight the drawbacks specific to demographic data as they prevent us from identifying, examining, and correcting systemic inequities.

⁷ Bedford Hills is one of several state-run prisons for women.

⁸ See [The Power and Problem of Criminal Justice Data: A 20 State Review](#).

Arduous and Costly Processes

Many of the people we interviewed described data-request processes that are rife with long wait times, resource depletion, and a dependence on associated costs. Both data requestors and data providers frequently named these processes as a hurdle to acquiring or providing meaningful information. While system administrators noted a tedious process is sometimes necessary to safeguard individual privacy in the data, an ordinary user who is not an expert is not in the position to navigate such procedures. The result: inconsistent access to information. At the local level, the difficulty and cost associated with obtaining data appears largely dependent on agency resourcing, creating wide variation in access across jurisdictions.

Reflecting on his time working with court data, New York Supreme Court Judge Craig Doran explained how accessing relatively basic information could take him days and likely exists in a format that would require additional effort on his part in order to make sense of. He explained that some of the challenges to data access are a “function of how [data are] kept and where and by whom, and sometimes it’s just a matter of logistics.” The logistical challenges associated with data sharing were cited by practitioners and researchers alike. One researcher noted that many Sheriff’s Offices struggle to extract data from their databases, often turning to system vendors or other third parties to run reports, which ultimately increases the time and cost involved with the request.

At the state level, DCJS’s Leigh Bates discussed what this experience is like from a data sharer’s perspective: “There are many steps,” she explained, as she outlined the process for receiving case-level data from DCJS:

“First, we conduct an internal review to determine if the project will likely benefit criminal justice practitioners in New York State and then we work with the researchers to ensure we understand the research question, sample, and project timeframes. Once that is complete, there is another review by our counsel’s office, often followed by final review and approval by agency leadership. It’s a pretty resource-intensive process to ensure that information is appropriately reviewed and provided.”

Carolyn Cadoret, Senior Information Technology Analyst with OCA, also alluded to these process difficulties when discussing the mere evaluation of requests received:

“We look at the purpose of the data request and examine how might it help improve our courts. Does the data requested match the needs and scope of the stated purpose? We also analyze the level of effort for our staff. Given that we have multiple case management systems, this evaluation process alone can require significant time, and that is just to determine how much time and effort it will ultimately take to complete the request.”

Bates discussed the importance of protecting confidentiality and how that can also impact the time involved with a data request, “For outside research requests, it’s a very comprehensive process but that is by design. It is our responsibility to protect the confidentiality and anonymity of individuals whose information is included in the data. Because of that, we’re very interested in making sure those checks and controls are in place and we’re giving information that is accurate, complete and responsible to the needs of the project.” Karen Kane with OCA echoed the importance of safeguarding sensitive information from unauthorized disclosure, “More and more we’re providing data for outside purposes. We are looking to balance providing the requested information while also ensuring that we’re protecting the privacy of the individual defendants.”

“ It was a really difficult process with such a long waiting time that made it almost impossible to investigate certain patterns. ”

While it’s critical to protect someone’s personally identifiable data, it’s also important to note that the time involved with evaluating a request often hinders opportunities to make policy changes based on relatively recent data. Julie Ciccolini, the Director of Law Enforcement Accountability at the National Association of Criminal Defense Lawyers, recalled requesting case-level data from DCJS while at her previous role at the Legal Aid Society, “It was a really difficult process with such a long waiting time that made it almost impossible to investigate certain patterns,” she explained, “If you wanted to figure something out and had a deadline of say, six months, that wouldn’t happen.”

Monroe County Public Defender Tim Donaher shared a comparable sentiment, “If I go to DCJS looking for data from today to six months ago, it’s not available...not until months later which is useless when I’m trying to make an argument about bail reform impact right now.” Similarly, researchers with CANY indicated that it can take two to three months to receive basic statistics from DOCCS, such as the Department’s “under custody reports”. These reports tell us essential information about incarcerated populations which can then be used by advocates to better understand the needs of the communities they’re serving. However, the more time that passes, the less relevant the information is to the current populations.

Several interviewees also commented that requests for information commonly involved an accompanying cost barrier and that agencies often provided cost estimates they felt were “irrational” and “excessive.”

“ They can take as long as they want and charge as much as they want. ”

One legal researcher reflected on her time as a defense attorney in New York City, recalling how difficult it was to obtain policing data and noting one department provided a cost estimate of \$47,000 to access misconduct information recently made available.⁹ “If people don’t want to track it they can throw out this exorbitant cost or estimate of the time it will take to pull the data,” said another researcher of their experience seeking data from local law enforcement. “They can take as long as they want and charge you as much as they want,” they said. While such an exorbitant cost certainly isn’t feasible for many data seekers, many of our interviewees cautioned against conflating genuine resourcing challenges with intentions to thwart transparency.

“ One department provided a cost estimate of \$47,000 to access misconduct information. ”

As was the case with data recording mechanisms, the challenges associated with data sharing are highly dependent on the resource capacity of a given agency, resulting in wide variation in data access between jurisdictions. For example, an individual might make the same request to two different county jails. While one may run a predefined report in their case management system and charge the requestor \$15 to pull the data, the other may require a week to compile paper records,

⁹ [NYS Senate Bill S8496](#) repealed section 50-a of the NYS Civil Rights Law which required the concealment of law enforcement disciplinary records.

charging \$600. Where data sharing processes have been made particularly laborious, practitioners are left without basic data necessary to their work. And not all requestors have the same ability to circumvent these barriers. Whether by way of legal knowledge, money, or position, some people are at an advantage, exacerbating disparate data access.

Access Inequity

Based on our interviews, it is clear that when it comes to criminal justice data in New York State, access appears to be granted to some and not others for reasons that are impenetrable at best, suspect at worst. For example, “They always approve my request,” and “We’ve tried for years to get access,” were statements we heard from two different individuals with reference to accessing the same type of data from OCA. Whether through a complete blockage or granting access to select pieces of information in particular formats, sentiments reflecting intentional barriers to data access were echoed across stakeholder types. Many of the people we spoke with highlighted the ways data access is granted haphazardly, even to the point of suggesting that some denials are an intentional obstruction.

What does New York State law say about data access?

Article 6, Sections 84-90, of the New York State Public Officers Law, also known as the [New York State Freedom of Information Law \(“NY-FOIL”\)](#), creates a presumption of access to the record of most criminal justice agencies in the state. The law covers government agencies at all levels, from state

agencies like DCJS and DOCCS, to District Attorneys at the county level and municipal policing agencies. The big exception is the courts. Court records are excluded from NY-FOIL and instead governed by Section 255 of the Judiciary Law. Additionally, the law emphasizes the importance of personal privacy and allows for the withholding of records containing personally identifiable details.

“The legal challenges are unbelievable, and they’re often just around interpretation.”

One avenue through which the preferential granting of access appears most frequently is by way of legal interpretation. Put simply, agencies, and oftentimes the individual people within those agencies, have different understandings about the laws governing data access. “The legal challenges are unbelievable and they’re often just around interpretation,” explained Iman Abid-Thompson as she described her experience requesting data in Monroe County. According to Abid-Thompson it is not uncommon to receive denials stating the information is not legally releasable. She wasn’t alone. Disparate interpretations of the rules were a commonly cited barrier across our interviews and several people shared the opinion that differences in interpretation have given agencies discretion over when, and to whom, they share data.

“The prison folks will tell you there’s a form, but in my 10 years I’ve never seen anyone get their information through that process.”

Michelle Lewin is an attorney and the Executive Director of the Parole Preparation Project, which provides advocacy and support to currently and formerly incarcerated people. She discussed her struggles attempting to access information from DOCCS and individual prison facilities. Many of her clients are serving life sentences with the possibility of release from prison through the state parole board. In order to prepare them for a fate-determining parole board interview, Lewin needs access to their prison and parole files, which the Board uses in their decisions. However, she has struggled to access these records, particularly the parole file, despite state regulation¹⁰ specifically permitting access to its contents. Moreover, her clients often struggle to receive their own legal documents, such as pre-sentencing reports, due to a lack of formalized infrastructure. While prison representatives direct her clients towards a form, Lewin said she has never had a client successfully obtain records via its filing.

Agencies also sometimes respond to similar requests with different data. MFJ experienced this firsthand when we sought court data and received a “media file.” Other researchers referenced receiving more comprehensive data extracts for their projects. Discrepancies like this were commonly cited, including instances where data are presented in a format that is difficult to

¹⁰ See NYCRR, Title 9, Subtitle CC, Section 8000.5.

work with - such as a paper file or scanned PDF - when electronic information is available and legally permissible or instances where the information released is pre-selected based on agency preference.

“ It was really just limiting the data to make it look like there was not really a problem. ”

Several people discussed these differences in interpretation in relation to requests for police disciplinary records. A 2020 law¹¹ repealed Section 50-a of the NYS Civil Rights Law which required the concealment of such data. Until then, this information was shielded from public scrutiny. “One department is saying the 50-a repeal was not retroactive, one saying it’s only retroactive,” one researcher explained, describing a statewide game of “tug of war” around the data. Relatedly, another person described a public data website created by one local law enforcement agency in response to the legislative change, noting it excluded certain disciplinary cases they had previously obtained through a NY-FOIL request. “It was really just limiting the data to make it look like there was not really a problem,” they explained.

Several stakeholders underscored how different access can be depending on who is making the request. One university professor explained that she has never encountered an issue accessing criminal justice data in New York State, crediting “great working relationships” and familiarity with the law. Additionally, because her work is well funded, cost has not been an issue for her in accessing data. However, these gateways do not exist for all people. Recall that researchers with CANY described waiting months for basic aggregate data on the state’s prison population. As Felicia Henry, CANY’s Director of Research and Policy, explained, “I think the climate becomes even more difficult with actors outside of the system. Even when it comes to us, it’s murky,” she said, referencing the independent organization’s legal authority to monitor DOCCS.¹² Another researcher agreed, “As an outsider, there are often charges that aren’t levied if you’re inside the government.” While interviewees commonly drew this distinction between those working “inside” and “outside” of the system when it comes to data access, not all practitioners on the inside are privy to the same information.

While many local practitioners described difficulties accessing information vital to their jobs, those working in the defense space faced a particular disadvantage as they are excluded from the state’s official eJustice portal, an integrated system designed to give qualified agencies access to current data on criminal cases throughout the state.

¹¹ See [NYS Senate Bill S8496](#).

¹² See [NYS Senate Bill S8046](#).

The list of “qualified agencies”¹³ permitted access to the system is long and includes courts, probation departments, sheriffs’ offices, district attorneys’ offices, state and local departments of correction and community supervision, and even the Onondaga County Center for Forensic Sciences Laboratory when acting within the scope of its law enforcement duties; but defense practitioners have been left off. Excluding defenders from the system leaves them at a clear disadvantage when providing client services. For example, in some instances, prior felony convictions serve as predicate offenses, enhancing the penalty for future felony convictions. Without access to a client’s criminal history data, defense providers may miss key information that can ultimately lead to a mandatory prison sentence.¹⁴ Moreover, with access to a client’s criminal history data, defense providers are in a better position to determine their client’s eligibility for various diversion programs and services.

“**It’s a fight we’ve been having for decades.**”

New York State Defenders Association Executive Director, Susan Bryant, discussed the struggle for access to that data, noting 2019 legislation that would have granted defenders access to the system, but was vetoed by then-Governor Andrew Cuomo. Bryant noted that a similar bill¹⁵ is currently under consideration that will expand the definition of qualified agencies to include public defenders, legal aid societies, and assigned counsel administrators. She is hopeful it will be signed this time around. “It’s a fight we’ve been having for decades,” she explained. Without access to the eJustice portal, indigent defense providers are left to pursue avenues for data collection available to the general public, resulting in the same challenges around interpretation and preferential access. One example highlighted in our interviews involves ongoing access to court data, which interviewees claim is granted to a single defense agency in the state for reasons unknown to others. “The other public defenders and myself have repeatedly requested that same data feed but have been unsuccessful so far,” explained Chris Boyle from New York County Defender Services.

¹³ See [NYS Executive Law Section 835](#) for the complete definition of qualified agencies.

¹⁴ See [NYS Penal Law Section 70.06](#), which requires the court to impose prison time for certain second felony convictions.

¹⁵ See [NYS Assembly Bill A7729](#).

Moving Forward:

Where do we go from here?

While this report has outlined a number of challenges relative to New York’s criminal justice data, it’s important to note a few things that are working well, particularly in New York City. Most of the people we spoke with drew a distinction between New York City and the rest of the state, noting the ways adequate resourcing can promote data successes at the local level. Initiatives like NYC OpenData, the result of a partnership between the Mayor’s Office of Data Analytics (MODA) and the Department of Technology and Telecommunications (DoITT), were referenced as examples of successful data sharing initiatives. Additionally, the New York City Criminal Justice Agency (CJA), in partnership with the New York City Mayor’s Office of Criminal Justice (MOCJ), recently published a Pretrial Release Dashboard,¹⁶ granting public access to data about people who were arrested and charged with an offense and subsequently released back into their community until their case’s disposition. This information is key to informing ongoing discussions about the impact of bail reform in the state. Researchers working with New York City data received from DCJS praised the consistent arraignment variables and low rate of missingness, attributing this to New York City’s advanced technological capabilities, such as robust case management software implemented by prosecutors that encourage a more comprehensive approach to data collection.

Moreover, the conversations underlying this report pointed to a heartening recognition of the importance of data among New York stakeholders, and a general interest in leveraging data to generate transparency, trust, and better decision-making. As one researcher explained, “Agencies want to be able to justify the work they’re doing and engage with policymakers around data, so I think there’s been a cultural shift to some extent.” The people we spoke with also described several routes that agencies looking to advance equitable and responsible data sharing might pursue.

¹⁶ See nycja.org/pretrial-release-dashboard to explore CJA’s Pretrial Release Dashboard.

Rethink the Purpose of Data Collection

When agencies refer to criminal justice data, they are most often referring to information they collect to support their daily operations. When researchers and advocates refer to data, they intend it for the purposes of assessing system performance. As a result, the data agencies often collect isn't always the data researchers and the public need. Many of the stakeholders we spoke with recognized the tension between these two competing purposes. Karen Kane, Director of Court Research with OCA, explained how the case management system's purpose is to keep an accurate court record and allow for the efficient processing of cases. "Therefore, there is a disconnect between what data the clerks need to do their job versus what researchers and policy advocates want or need for data analysis," she explained, "The case management system was not designed or intended to be a research database."

This disconnect means we're often left without the information necessary to measure system performance or examine the impact of particular policies. In order to achieve these broader goals, New York stakeholders should consider rethinking the purpose of data collection and consider ways to more accurately account for peoples' experience with the system. This shift can be done in a number of ways, from beginning to record self-reported race to implementing a plan to follow up with defendants about their experiences in court.

Establish Data Standards

Criminal justice data recording practices across New York State look wildly different by jurisdiction, agency type, and even the individual entering information into a case management system on a given day. Because of these differences, even when people uncover a pathway to data access, it remains challenging to compare practices from one county to another, making

statewide analyses often impossible. While the stories shared here tell us New York State has a way to go in this regard, several people cited ongoing efforts to create agreed upon terms and definitions. For example, people working in the prosecution space described efforts led by the state's Justice Task Force¹⁷ to unify and consolidate disposition codes so we can accurately see how different cases conclude. Only with parallel, consistent information can District Attorneys better understand their practices relative to others in the state.

Moreover, the Manhattan District Attorney's Office was cited by multiple people as a leader in establishing accurate and consistent data recording practices that continue to evolve over time. Still, when it comes to standardizing data, there is room for progress at every level. For example, New York State stakeholders might consider pursuing statewide legislation that mandates the recording of clearly defined data elements. Indeed, this need was commonly cited by the people we interviewed as the most promising pathway to change. Beyond legislation, local agency leaders can continue to make data standardization a priority by way of collaborating with one another around definitions and collection practices. Even within individual agencies, there is ample room to standardize the ways in which we define and record data and, as a result, improve the overall quality of the information collected.

Address Ad Hoc Transparency

Even if data are well-defined and recorded consistently, at the end of the day people cannot request access to information they aren't aware exists. The ability to obtain and analyze data should not be based on "who you know" or limited to those privy to information about an office's practices or records management system. Yet this is a running theme with those we interviewed. Mandating the collection of certain data elements has the potential to address this inequity to some extent, by way of communicating the information agencies are required to track. The new OCA-STAT¹⁸ Act Dashboard, created by the Division of Technology and Court Research (DoTCR), is one example of the ways legislation can spur impartial transparency. Through publicly accessible tables and graphs, the dashboard presents aggregate data on low-level offenses as well as felonies. The dashboard is paired with a downloadable extract for researchers to explore further trends and includes a comprehensive codebook with field definitions.

¹⁷ A permanent task force established in 2009 to recommend changes that will increase the fairness and effectiveness of the NYS criminal justice system

¹⁸ Judiciary Law § 212 (2) requires the UCS to compile and release data on low-level offenses

However, agencies throughout New York State that value transparency needn't wait for a mandate to begin making accurate information about the data they record publicly available. DCJS lends one example. In addition to information statutorily required to be posted, the agency posts a variety of different crime, arrest, jail and disposition data sets in downloadable formats on its public website.¹⁹ Removing the obscurity around what information agencies collect is necessary for equitable data sharing practices and an essential step towards true transparency.

Alleviate Capacity Issues

As noted throughout this report, agency capacity can have a profound impact on the quality of data collected, as well as how accessible it is made. The data providers we spoke with described successes they credited to adequate funding, as well as data aspirations that have been hindered by a lack of staff or technology. As one District Attorney's office in Western New York put it, "If you're going to talk about data, you have to talk about money. No use in building this pie-in-the-sky that no one can afford." This sentiment was echoed by those working across different segments of the criminal justice system and, after legal mandates, resourcing was cited as the most influential factor in shaping the quality and availability of data in a given place. When agencies are under-resourced, data tracking and dissemination often take a back seat to duties thought of as more essential. While there is no simple or "one-size-fit-all" solution to this constraint, the burden will likely decrease over time as agencies take steps toward standardization.

For example, determining standardized formats for data was cited by agencies at the state and local levels as a way to reduce the burdens of customized requests. Other practitioners discussed how collaborating and exchanging data with other agencies within their jurisdiction has streamlined data processes and reduced associated costs. Whichever path to better data New York stakeholders pursue, recognizing that agencies are working at vastly different capacities, and ensuring they have appropriate resourcing allocated to data objectives will be key.

¹⁹ See <https://www.criminaljustice.ny.gov/crimnet/ojsa/stats.htm> for more information.

Conclusion:

Data are an incredibly important tool in helping to shape our understanding of the criminal justice system. With data we have the power to uncover trends in the way cases are processed through the system, hone in on problems, and identify effective solutions. The extent to which stakeholders in New York State have difficulty obtaining data they need, either due to poor quality or a lack of accessibility, suggests that a lot of what unfolds in criminal justice agencies across the state goes unseen. This opacity prevents us from engaging in evidence-based discussions related to criminal justice policies and practices. As Katie Schaffer, Director of Organizing and Advocacy at the Center for Community Alternatives, explained, “It means we often know anecdotally there are problems, but when we go to find the data necessary to assess it at the systems level, the data are obfuscated.”

While it is difficult to separate the factors influencing decisions to make data publicly accessible, the stakeholders we spoke with indicated that, at the least, perceptions of obstruction are impacting trust in the system. Moreover, without data that can speak to what is happening, or information about what data are recorded, claims about the performance of the system go unchecked. Data means power and those with data ultimately maintain control of the system itself. When data are harbored in the hands of a select few, those without access are left at a disadvantage in ensuring the narratives told about crime and justice in their communities are a reflection of reality.

In addition to shielding the system from scrutiny, many practitioners, particularly those working on behalf of clients entangled in the system, are left without the information necessary to effectively execute their responsibilities. Withholding critical case information carries the potential to impact individual case outcomes. When police, prosecution, and the courts have access to timely criminal history information that is withheld from the defense, key information that can support their client’s case goes uncovered. Of particular concern were those examples cited where people who have been charged with, or convicted of, a crime and were denied access to their own records.

While the issues described here are troubling, there is reason to be hopeful. This report comes during a time of great opportunity as states across the country turn to understanding and improving their criminal justice data infrastructure. Indeed, our interviews showed stakeholders across New York State are no exception to this heightened data awareness. Beyond recognizing the power of data as a gateway to informed decision-making, many of the people we spoke with pointed to data as key to building trust between communities and the systems that ostensibly seek to deliver safety and justice. As one judge put it: “When the community you serve doesn’t trust you one might ask, what do you have to lose? You can only get better.” Criminal justice agencies across New York State should not lose sight of the power and promise of transparency as they consider new initiatives around the data they’re tracking, as well as who they share it with.

MFJ Contributors:

Shelby Davis | Assistant Director of Research and Strategic Initiatives

Jennie Brooks | Assistant Director of Data Outreach

Fiona Maazel | Director of Communications

With special thanks to the Tow Foundation for funding this work.

