WASHINGTON STATE STATISTICAL ANALYSIS CENTER

Criminal Justice Research & Statistics Center

Informing a data-driven justice system

The Justice Data Warehouse (JDW): The Data Handbook

Vasiliki Georgoulas-Sherry, Ph.D. & Hanna Hernandez, M.A.

Contents

Abstract

Washington state's criminal justice system has long operated in disconnected silos across federal, state, local, and Tribal levels, limiting the ability to assess performance, address disparities, and improve outcomes.

To respond to these impacts, the Washington Statistical Analysis Center (SAC) applied for and received the 2021 State Justice Statistics (SJS) grant from the Bureau of Justice Statistics (BJS). To tackle these challenges, the SAC, in partnership with the Public Safety Policy and Research Center (PSPRC), established the Justice Data Warehouse (JDW). This integrated platform links data from courts, jails, prisons, community supervision, and more, offering a comprehensive, longitudinal view of individuals' justice system involvement. The JDW enhances transparency, supports data-driven policy, and enables cross-sector collaboration with behavioral health, housing, and social services. It empowers interested parties to identify system gaps, evaluate reforms, and design more equitable, effective interventions. By centralizing justice data and supporting Tribal and local jurisdictions, the JDW plays a pivotal role in building a smarter, fairer, and more accountable justice system for Washington state.

Background

Like other criminal justice and legal systems in the nation, Washington state is insular; each decision point of the criminal justice and legal system is sectioned into silos, which includes federal, state, local, and Tribal agencies. Efforts continue to be made to shift from this siloed approach to a more integrated system.

Similar to other states, Washington lacks data integration capabilities that would allow for a comprehensive assessment of the criminal justice and legal systems' performance and outcomes. This is true at the local and Tribal levels, as well as the state level. For example, measuring performance and outcomes across systems (such as jails, courts and juvenile detention, community supervision, and prison incarceration) are hindered by a lack of integration. Addressing the need for integration and promoting collaboration across different components of the criminal justice and legal system can help bridge these silos and create a more effective system.

With this in mind, the SAC, in conjunction with the PSPRC, created the JDW to integrate criminal justice databases together as currently, each agency maintains their own distinct data sets. The JDW represents a significant step toward centralizing disparate data systems and enabling a holistic view of Washington's justice system. By linking data across agencies and jurisdictions, the warehouse supports a unified approach to understanding how individuals move through the system and where interventions may be most needed. For example, the ability to track an individual's journey from arrest through court proceedings, incarceration, and community supervision provides insight into system bottlenecks, disparities in treatment or outcomes, and the effectiveness of rehabilitative programs. This capacity for longitudinal analysis is critical for developing policies rooted in evidence rather than assumptions.

Moreover, the JDW strengthens transparency and accountability. Without integrated data, it is nearly impossible for policymakers, researchers, and the public to access meaningful information about systemic issues such as racial disparities, recidivism, or case processing times. With standardized and accessible data, interested parties can assess the impact of laws and policies in real time, uncover inequities, and identify areas in need of reform. In particular, Tribal and local jurisdictions — which often lack the technical infrastructure or resources to conduct their own large-scale evaluations — benefit from a shared platform that can illuminate how their populations are affected by broader justice trends.

The warehouse also fosters cross-sector collaboration, bridging the gap between criminal justice, behavioral health, social services, and public safety. Many individuals who enter the justice system do so at the intersection of complex issues such as poverty, mental illness, substance use, and housing instability. By integrating justice data with health and social service datasets, the JDW enables agencies to design and coordinate more effective, human-centered interventions. This holistic view not only enhances service delivery but also aids in identifying and addressing systemic drivers of justice involvement, especially among youth, people of color, and those with untreated behavioral health needs. The JDW enables interested parties to see how individuals interact with multiple components of the justice system — from arrest to reentry — creating a full picture of justice involvement.

Using Integrated Justice Data

Using integrated administrative data from criminal justice agencies can harness information in meaningful ways that transcend traditional "silos" and allows communities to focus collective attention on important social issues that cross systemic boundaries. As individuals in the criminal justice system frequently interact with other decision points (e.g., policing/arrests, trial/sentencing, and incarceration/community supervision), collecting and analyzing data across multiple criminal justice systems is critical to better understand the impact of the system. As the central cross-sector repository for Washington state's criminal justice and legal systems, the JDW can be used to house information from federal, state, local, and Tribal agencies, which can consolidate individual- and system-level data.

The JDW was developed to integrate siloed criminal justice data and can be used to support innovative and rigorous research methods that can provide answers to basic research questions as well as practical, applied solutions. Using quantitative tools like the JDW to influence qualitative decision-making, the JDW can tell the story of the people served across this state and the individuals who serve them within state government. This includes developing metrics, monitoring trends, disseminating reports, and delivering presentations to educate state agencies, interested parties, and impacted communities.

Through the JDW, data can be translated into meaningful information that drives the state's commitment to achieve equitable and just outcomes for Washingtonians. The JDW can be utilized to support legislative requests, program evaluations, operational decisions, and in-depth research. Furthermore, the JDW can help increase the transparency of the criminal justice and legal system to those outside of the system, such as court users, advocates, policymakers and lawmakers, researchers, and all Washingtonians.

The JDW's integration would also allow us to pinpoint data to particular areas of intervention (e.g., overrepresentation of person of color at all stages of the criminal/legal system, disparities in outcomes for marginalized individuals who interact with the criminal legal system, etc.). Additionally, the JDW could also be used to track and improve fairness in areas such as prosecution, adjudication, disposition, sentencing, incarceration, release, and community supervision.

An integrated justice data system like the JDW is essential for addressing the limitations of fragmented data that hinder effective system-wide understanding and response. Without integration, agencies often operate in isolation, making it difficult to track an individual's full trajectory through the criminal justice system or assess the cumulative impact of decisions made at different stages. Integrated data allows interested parties to break down these silos and view the system as an interconnected whole. This holistic perspective is critical for identifying gaps in service delivery, or unintended consequences of policy decisions. It also ensures that reforms in one part of the system — such as changes in sentencing guidelines — can be evaluated in terms of their ripple effects throughout the rest of the system.

The benefits of integrated justice data extend beyond operational efficiency and into improved outcomes for individuals and communities. By enabling comprehensive analysis, the JDW helps illuminate patterns and trends that may otherwise remain hidden — such as racial and geographic disparities, repeat justice involvement, or systemic delays in case processing. These insights support data-informed decision-making that can lead to more effective, equitable, and transparent justice policies. Additionally, integrated data empowers community partners, advocates, and researchers to collaborate on solutions grounded in evidence. Ultimately, a centralized data system enhances accountability, supports smarter resource allocation, and strengthens the state's ability to deliver fairer, more responsive justice to all Washingtonians.

Cross-Sector Research and Analysis in Justice Data

By fostering collaboration among criminal justice agencies, public health organizations, social services, and research institutions, the JDW will ensure that justice system data is not only comprehensive but also actionable. A well-integrated repository allows interested parties to analyze trends across law enforcement, courts, corrections, and community supervision, helping to identify systemic gaps, high-risk populations, and opportunities for reform. This level of coordination enables policymakers to implement targeted, data-driven interventions — such as diversion programs, sentencing alternatives, and reentry support — that promote equity, efficiency, and community well-being.

The integration of diverse data sources — including arrest records, court filings, jail and prison data, behavioral health information, and recidivism rates — provides a more complete picture of the justice system's functionality and impact. This comprehensive dataset empowers researchers to examine how factors such as mental health, substance use, socioeconomic status, and race or ethnicity influence justice involvement and outcomes. Additionally, the ability to track individuals' experiences across the justice continuum over time supports the identification of disparities, ensuring that vulnerable or marginalized groups receive the attention and resources needed to achieve fairer and more just outcomes. The JDW will serve as a critical tool for guiding evidence-based policy decisions that center on justice reform, equity, and effectiveness.

By leveraging cross-sector collaboration, the JDW will also support the development of proactive strategies that prevent justice involvement and reduce recidivism. With access to high-quality, integrated data, researchers and policymakers can identify patterns and emerging risks, enabling early interventions such as pretrial support, restorative justice practices, and mental health or substance use treatment programs. These preventative strategies not only reduce system burden and incarceration rates but also contribute to safer, healthier communities across Washington. Ultimately, the JDW's comprehensive approach to justice data analysis will play a vital role in shaping policies that reduce harm, improve outcomes, and enhance public trust in the justice system.

Cross-sector research and analyses are essential for improving justice outcomes and public safety by fostering collaboration between criminal justice, behavioral health, housing, education, and policy sectors. Justice system involvement is often linked to broader social determinants, and addressing these issues requires an integrated approach that considers root causes, systemic inequities, and community-based solutions. By combining data from multiple agencies and service providers, experts can identify key intervention points, common pathways to system involvement, and the most effective strategies for rehabilitation and support. This data-driven approach helps develop alternatives to incarceration, reduce recidivism, and address the long-term impacts of justice involvement on individuals and communities.

Technological advancements and innovative justice policies also benefit significantly from cross-sector

collaboration. For example, partnerships between corrections, public health, and social services have led to advancements such as real-time data dashboards, improved case management tools, and coordinated care networks for justice-involved individuals. Research that integrates data from courts, law enforcement, and community health providers can highlight the impact of pretrial detention, sentencing practices, or reentry barriers on long-term outcomes. By analyzing this information, policymakers can implement targeted reforms, community-based supports, and policies that reduce incarceration while improving public safety and social equity.

Cross-sector analyses also inform long-term strategies that promote restorative justice, reduce over-reliance on incarceration, and strengthen community resilience. Public health professionals and justice system leaders can advocate for initiatives that address trauma, reduce youth system involvement, and expand access to education, employment, and housing for those impacted by the system. Economic research further supports these efforts by demonstrating the fiscal benefits of investing in upstream prevention, community supports, and evidence-based alternatives to incarceration. Through cross-sector collaboration, Washington can build a justice system that is data-informed, people-centered, and committed to equity and accountability.

JDW Core Data

The SAC, with the support of the PSPRC, receives a variety of administrative datasets from agency partners that are incorporated into the JDW. These administrative datasets are outlined in Table 1, based on the category of data and data source – Administrative Office of the Courts (AOC), Department of Corrections (DOC), Caseload Forecasting (CFC), Washington Association of Sheriffs & Police Chiefs (WASPC)'s Jail Booking Reporting System (JBRS) and Washington State Patrol (WSP)'s Computer Criminal History (CCH) data. To supplement Table 1, Appendix A shows the data tables that are integrated into the JDW by agency, table name and purpose, years covered, and number of data elements. It also shows the data elements of the JDW by agency, table name, variables name, definition, and data type. These datasets vary in subject matter, from the initial police interaction to the arrest, to the booking, to court interaction, to sentencing, to incarceration and community supervision. As such, the JDW is one of the most comprehensive longitudinal criminal justice data systems in the state.

Table 1: The JDW Core Data Sources

Type of Data	Agency	Description
Carceral and Community Supervision	Washington State Department of Corrections (DOC)	DOC maintains information for people incarcerated in Washington state correctional facilities and for people under community supervision in Washington state.
Court Case Filings – Judicial Information System	Washington State Administrative Office of the Courts (AOC)	AOC maintains statewide electronic court records database for all cases seen by courts in Washington state.
CFC Convictions	Caseload Forecast Council (CFC)	CFC maintains the adult conviction database including data related to the crime, the offender, the sentencing judge, the sentence, and alternatives to incarceration.
Jail Booking and Reporting System (JBRS)	Washington Association of Sheriffs & Police Chiefs (WASPC)	JBRS (RCW 36.28A.040) is a multi-jurisdictional database providing criminal justice agencies an instant, up-to-date database of booking and release records from all city and county jails in Washington state.
Computerized Criminal History (CCH)	Washington State Patrol (WSP)	The WSP maintains a database of Washington criminal history information, or background checks, consisting of fingerprint-based records and disposition information from law enforcement agencies and courts throughout the state.

JDW Administrative Data Limitations

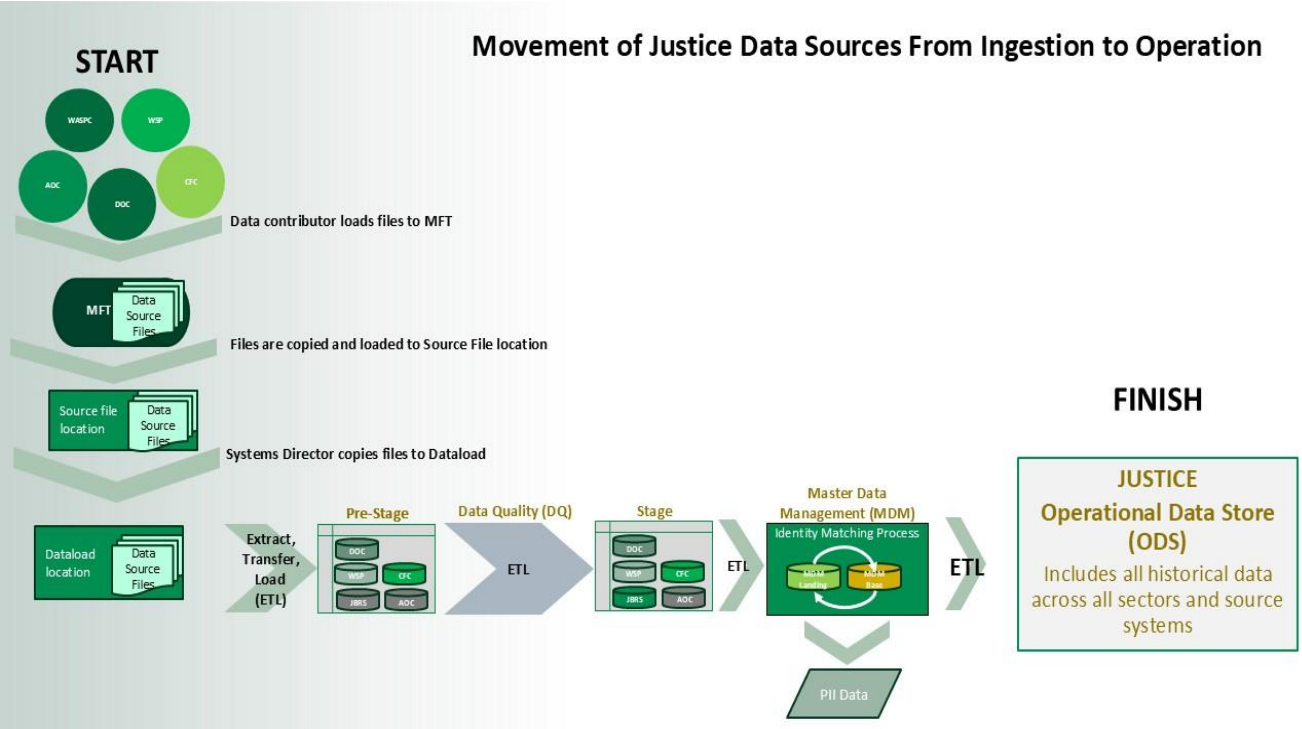
While all the datasets above are processed to the highest quality standards by the data contributing agencies, it is important to recognize that inaccuracies may exist within administrative data. Unlike other data, where both cross- and within-subject controls are possible, such measures are often unfeasible and impossible to incorporate in administrative data. Administrative data is also not typically collected for research or evaluation purposes but to meet the administrative needs of specific programs and specific state or federal reporting or monitoring requirements. Administrative data is collected as both transactional and summative datasets by local administrators and submitted to an agency authority, making variance among data collectors a potential source of bias in each dataset. Quality control processes may be imposed after data is submitted to agency authorities, which could impact data quality in ways that are difficult to detect within the final dataset.

The limitations are not meant to suggest that the administrative data loaded into the JDW is unreliable but rather to advise researchers to keep these potential concerns in mind as they request data and conduct research. Administrative data must always be thought of as the combination of both the collected data and the process used to collect the data. The data summaries in this Handbook delve into these processes. Researchers who use the JDW for analysis purposes should review all the available data documentation and adjust their models according to the research question and the administrative data collection procedures.

Flow of Contributor Data

Figure 1 illustrates how data from contributing agencies is loaded into the JDW. Once data is received through a secure file transfer process, it is loaded to a pre-stage database, then undergoes a series of quality checks before it is transferred to a stage database.

Figure 1: Flowchart of Movement of Justice Data Sources From Ingestion to Operation



Personally identifiable information (PII) is separated at that point from the rest of the data and used for identity resolution. Once the identity resolution process is complete, de-identified data are moved into the JDW Operational Data Store (ODS) and become available for analysis. All PII data from the source files are excluded from the JDW data repository and is unavailable to researchers. JDW Analytical Files include selected de-identified data elements from the JDW ODS that represent several public safety outcomes and measures. The JDW was created by extracting, transforming, and loading the data using the minimum necessary combination of variables needed to uniquely identify a person in a specific dataset dependent on what is available. Figure 1 illustrates the data flow and loading process for the JDW. The process includes: (1) receiving initial data profile (e.g., did we get what we were expecting, counts); (2) updating ETL/crosswalks, if necessary (e.g., new fields, new codes); (3) loading to STG, verify (e.g., Powercenter, integrated error flags); and (4) loading to MDM, ODS (e.g., Powercenter). The source data is rightsized, and the loading function is then centralized.

The JDW linkage process has two key parts: (1) the creation of a Primary Key (PKey) and (2) a Linkage ID (ID). A PKey is the minimum necessary combination of variables needed to uniquely identify a person in a specific data set. The minimum necessary parts of a PKey will depend on what is available in any given data set. A complete PKey will have all the minimum variables required attached to one record. An example of a complete PKey would be someone's first name, middle name, last name, and date of birth having at least a first name, last name, and date of birth.

One way to increase the possibility of improving the matching results is having a tiered PKey system, which requires more information in the data set to attempt. To follow on the example previously used, if a data set were to also have the last four digits of a social security number (SSN), it may be possible to have the first set of PKeys using a person's full name, date of birth, and the last four of the SSN. If the first set of PKeys is incomplete, the second set of PKeys could be a person's full name and date of birth. Identifying and creating a complete PKey is the first step in the linkage process.

Once a complete PKey has been identified and created, an ID is bound to all records with those associated PKeys. The relationship between the PKey and the ID is a "many-to-one" relationship. The linkage process uses both probabilistic and a deterministic matching algorithm. The two matching methods leverage all available data sets in the OFM data warehouse. As the data warehouse expands, the linkage outcomes will change. A series of quality control checks to improve the linkage results are conducted once the matching is done. The linkage checks start with a set of automated rules leading to an as-needed manual review of remaining issues with matching.

Identity matching is conducted through the utilization of Informatica, which provides a central repository of identifiers (e.g., full name, date of birth, and, when available, SSN) over time for each individual source. Data that populates the JDW comes to OFM from each individual data provider in various formats.

Data then enters the JDW Pre-Stage, in which data quality is assessed. At this time, differences between current and previous data sets may be discovered and addressed. Next, data moves to the JDW Stage, where identifiers are cleaned and standardized. For each record, a standard set of identifiers are strung together to create an identity resolution token. These tokens are constructed so that no two people will share the same token. In addition, names are standardized using a rigorous set of rules incorporated into Informatica.

Data then moves to Master Data Management (MDM), where the new data and tokens are merged with existing data already in the JDW. For example, if a person already exists in the data warehouse, their new data (tokens) will be assigned to the existing ID. Once the identity matching is complete and all data

has been merged, it is moved into the JDW ODS where it can be used for research projects, data sets, etc. Data in the ODS is public facing and does not include PII.

Identity Resolution Process

The core feature of the JDW is the linking of cross-sector data. Through an identity resolution process, JDW links individuals across data files from contributing agencies to facilitate longitudinal and cross-sector analysis. Identity resolution is the process of identifying records that belong to the same entity (e.g. person or household). The purpose of JDW's identity resolution process is to identify and create linkages across multiple data sources so that crash records associated with a given individual are linked to related records associated with that individual and event in the roadway, police interaction, court interaction, and health encounters datasets. For the JDW repository, this involves linking individual-level data, such as names and birth dates, across multiple sources and identifying these individuals with unique person identifiers. These identifiers are referred to as "PersonIDs." PersonIDs are assigned to all individual-level data received by JDW from our data contributors. As additional linking activities occur, PersonIDs are updated to reflect the most recent data available.

It is important to understand the identity resolution process so the researcher can evaluate whether it may impact the analysis, especially if the research includes linking the JDW data to additional data.

Creation of Token IDs and Assignment of PersonIDs

Before individual-level PersonIDs can be created or assigned, identity resolution tokens referred to as a "PKeys" are created for each record in a dataset. A "PKey" is an identifier or combination of identifiers from a dataset that are unique to an individual within the dataset. JDW reviews each dataset to establish PKeys specific to the dataset. Identifier and individual characteristic data from the datasets and associated with the PKeys are then loaded into the identity resolution system for linkage. PKeys that already exist in the identity resolution system are attributed with the PersonID assigned to that PKey. PKeys that do not already exist in the identity resolution system are assigned a unique preliminary PersonID which may be overwritten in the identity resolution steps described below.

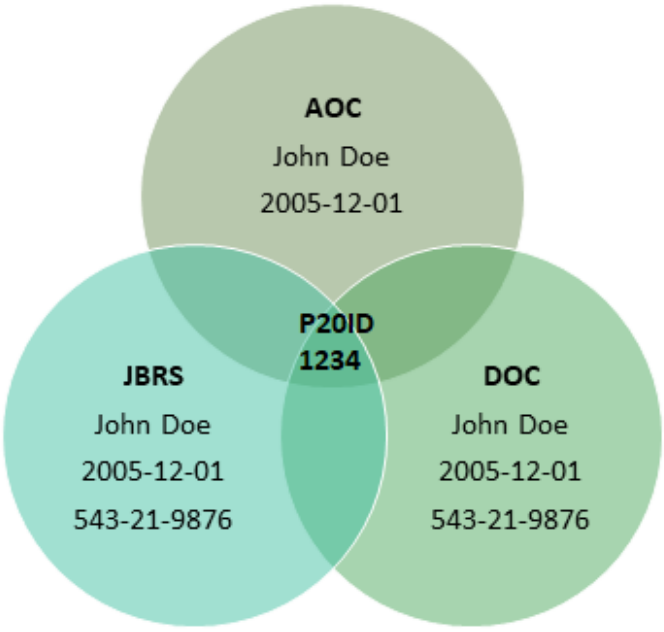
No single set of identifiers is common to all data sources, so the identity resolution process and match rules are tailored to the source of data being matched. For example, DOH data has names, birth date, and Social Security Number (SSN) whereas AOC data has names and birth date. As a result, rules linking each are limited to the available identifiers (see Table 2).

Table 2: Identity Token Components in Data Sources

Sector	Agency	Birthdate	First Name	Middle Name	Last Name	SSN
AOC	AOC	**X**	**x**	**x**	**x**	
CFC	CFC	**X**	**x**	**x**	**x**	
DOC	DOC	**X**	**x**	**x**	**x**	**x**
WASPC	JBRS	**X**	**x**	**x**	**x**	**x**
WSP	CCH	**X**	**x**	**x**	**x**	**x**

Two datasets with different sets of available identifiers can be indirectly matched by involving other sources of individual data. For example, the Licensing dataset above can only be directly linked to the Collision data, but the Collision data can be linked to all other datasets. Consequently, the licensing data may be linked *indirectly* to all datasets in the JDW repository (see Figure 2).

Figure 2: Identity Token Components in Data Sources



Phases of Identity Resolution

JDW's identity resolution process has four phases which are performed sequentially for each iteration of linkage.

1. Blocking
2. Evaluation
3. Cardinality analysis
4. Merging

Blocking

Blocking is the process of creating potential pairs between new PersonIDs and pre-existing PersonIDs in the identity resolution system. This is accomplished by a set of match rules. An example of a match rule is "match two records if they have the same names, same birth dates, and same SSNs." When new PersonIDs are matched against the pre-existing PersonIDs, the result is a set of prospective PersonID match pairs.

The quality of the match pairs depends on the match rule. For match rules involving the exact match between many ID fields, such as the "same names, same birthdates, and same SSNs", the false positive rate is low. The match quality is lower when the match rules involve a small number of fields, when fuzzy logic (such as the SOUNDEX algorithm[1]) is used, or when matching on partial strings within name fields.

Once blocking is done, the identity resolution system calculates a probabilistic match score for each potential match pair using the Expectation Maximization (EM) algorithm.

[1] SOUNDEX is an algorithm that indexes words based on their sound. This enables comparisons of words or names based on phonetic differences. https://en.wikipedia.org/wiki/Soundex

Evaluation

The set of potential match pairs is split into three categories:

- **High-probability matches.** These PersonID match pairs are the result of application of conservative match rules (e.g., "same name, same birth date, same SSN"). For this set, undermatching (or not correctly identifying an actual match) is not a significant concern, as the conservative match rules are designed to ensure extremely low false positive rates. Probabilistic match scores might also be used to further delineate the set of high probability matches.
- **Mid-probability matches.** These potential match pairs result from the application of looser match rules than the ones used to create the high probability matches (e.g., "same name, same birth date"). The match pairs and probabilistic scores, and associated identifiers (names, dates of birth, etc.) are brought into one dataset. The match pairs are then manually reviewed in Excel. The match pairs that are deemed to be actual positive match pairs are flagged, and the results are integrated into the prospective match pairs within the identity resolution system. At this point, the identity resolution system contains the set of provisional match pairs.
- **Low-probability matches.** These potential match pairs result from very loose match rules (e.g., "same county of residence, same gender, same first name"). Based on low probabilistic match scores, none or very few of these potential match pairs are provisional match pairs, hence these potential match pairs are ignored.

Cardinality Analysis

Cardinality analysis is a key component of the identity resolution process. It allows for more aggressive matching, while at the same time, improving the quality of the existing PersonID linkages in identity resolution system. In cardinality analysis, the provisional match pairs are merged on a trial basis. Then the cardinal relationships are determined between the PersonIDs in the subject dataset and those in the repository being matched against. These relationships can be 1:1, 1:Many, Many:1, or Many:Many. For example, a 1:Many relationship indicates that one PersonID in the subject dataset matches multiple PersonIDs in the universe of data being matched against. The 1:1 relationship, where two PersonIDs are paired exclusively, are accepted. The relationships involving non-1:1 relationship are manually reviewed to resolve them as accurately as possible.

In addition to reviewing matches of records for the current dataset, the cardinality analysis step permits the analyst the opportunity to review matched records from previous iterations of the identity resolution process. The analyst may choose to merge or unmerge previously matched pairs based on new information associated with the records. Consequently, the identity resolution process yields continually improving linkages across datasets.

Once the cardinal relationships are verified, the results are fed back into the identity resolution system. The result is that some PersonIDs might be unmerged, some could be merged, and some provisional match pairs might be deemed not to be positive match pairs.

Merging

After the cardinality phase is concluded, the match table now contains a list of positive match pairs of PersonIDs. These match pairs are then incorporated into the identity resolution system using an automated process. The result is that people who had been previously represented by multiple preliminary PersonIDs are now represented by a single PersonID.

a positive match. The probabilistic score, match pairs, and associated identifiers (names, dates of birth, etc.) are brought into the same dataset. The match pairs are then manually reviewed, and the match pairs that are deemed to be actual positive match pairs are flagged. The results are integrated into the prospective match pairs within MDM. At this point, MDM contains the set of provisional match pairs. These prospective match pairs are then subject to cardinality analysis.

Cardinality Analysis Phase

Cardinality analysis is a key component of the identity resolution process. It allows for more aggressive matching, while at the same time, improving the quality of the existing PersonIDs in MDM. In cardinality analysis, the provisional match pairs are merged on a trial basis. Then the cardinal relationships are determined between the PersonIDs in the subject dataset and those in the repository being matched against. These relationships can be 1:1, 1:Many, Many:1, or Many:Many. For example, a 1:Many relationships indicates that one PersonID in the subject dataset matches multiple PersonIDs in the universe of data being matched against. The 1:1 relationship, where PersonID matches exactly, are accepted. The relationships involving non-1:1 relationship need to be manually reviewed to resolve them as accurately as possible.

Once the cardinal relationships are verified, the results are fed back into MDM. The result is that some PersonIDs might be unmerged, some could be merged, and some provisional match pairs potentially are shown not to be actual match pairs. At the conclusion of this iterative process, the PersonIDs are integrated into the data contained in the ODS for use in analysis.

Core JDW Contributor Data

This section provides a set of descriptions or quick references to the core data files that feed into the JDW . This information should be used to:

- Guide researchers toward data that are relevant to their research questions.
- Provide metadata that will inform research design.
- Provide examples of how the data is used in research.

Washington State Patrol

The Washington State Patrol (WSP) Computerized Criminal History (CCH) system is the state's central repository for criminal history record information. Managed by the WSP's Criminal Records Division, the CCH contains detailed information on individuals who have been arrested and fingerprinted in Washington. This includes data on arrests, charges, court dispositions, sentencing outcomes, and correctional supervision. As such, the CCH serves as a foundational tool for background checks, law enforcement investigations, and public safety decision-making across multiple sectors.

CCH data is used extensively by law enforcement, the courts, licensing agencies, and employers for purposes ranging from criminal investigations to firearm purchases, child welfare screening, and employment eligibility. The WSP CCH system is also integrated with the national Interstate Identification Index (III), allowing criminal history information to be shared across states. Importantly, only arrests that result in fingerprinting are included, ensuring a standardized and verifiable record. Dispositions are supposed to be submitted by the courts to ensure completeness, but not all agencies consistently report outcomes.

Despite its value, the CCH system has limitations that can affect both accuracy and usability. Incomplete disposition reporting is a persistent challenge, often leaving arrest records without corresponding court outcomes. This gap can result in misleading or outdated information being used in critical decision-making. Additionally, delays in data entry, inconsistent reporting practices, and technological disparities across local jurisdictions contribute to data fragmentation. These limitations highlight the need for improved data-sharing agreements, automation, and uniform reporting practices across the justice system.

Integrating CCH data into broader criminal justice data systems like the JDW offers significant opportunities. Doing so would enable more comprehensive tracking of justice-involved individuals across the entire system — from arrest to adjudication and beyond. Researchers and policymakers could better analyze trends in arrest rates, charging practices, recidivism, and sentencing outcomes. It would also provide critical insights into disparities by race, geography, and socioeconomic status, supporting efforts to advance equity and transparency in the justice system.

In the future, strengthening the CCH system through cross-agency collaboration, improved data governance, and modernized infrastructure will be key. Automating disposition reporting from courts, ensuring real-time data updates, and linking CCH data with other datasets — such as behavioral health or housing — could support a more holistic understanding of criminal justice involvement. As Washington state moves toward greater integration and data-informed decision-making, WSP's CCH system will remain a critical asset for improving accountability, guiding reform, and enhancing public safety.

Washington Association of Sheriffs and Police Chiefs

The Washington Association of Sheriffs and Police Chiefs (WASPC)'s Jail Booking Reporting System (JBRS) is a statewide data collection effort that gathers standardized information on individuals booked into jails across Washington. This system captures critical data points including the date and time of booking, charges filed, demographics (such as age, gender, and race/ethnicity), arresting agency, bail or bond amounts, and release dates and conditions. JBRS provides a valuable snapshot of who is entering local detention facilities, for what reasons, and under what legal circumstances. This real-time data is crucial for understanding patterns in pretrial detention and identifying disparities in jail populations.

One of the primary benefits of JBRS data is its ability to support policy decisions aimed at reducing unnecessary incarceration and improving pretrial justice. By analyzing trends in bookings — such as the proportion of individuals held on low-level or nonviolent offenses, or those unable to pay bail — policymakers and interested parties within the justice system can implement reforms that prioritize release over detention for eligible individuals. JBRS also provides insights into law enforcement and court practices across jurisdictions, allowing for evaluation of how charging decisions and arrest trends may vary by region, population, or time period.

Despite its utility, JBRS data face limitations related to consistency, participation, and completeness. Not all counties or jail facilities report data with the same level of detail or frequency, and some may lack the technological capacity to automate data entry and sharing. This results in gaps that complicate statewide analysis and hinder efforts to identify broader systemic trends. Additionally, while JBRS captures initial booking data, it may not always link to court outcomes or correctional data, making it difficult to track what happens to individuals after their initial jail stay.

The integration of JBRS data with other justice and social service datasets — such as court proceedings, behavioral health records, or community supervision outcomes — can vastly improve its impact. For example, linking JBRS to court data allows interested parties to analyze case resolution timelines or pretrial release decisions, while connections to health data can reveal the prevalence of mental illness or substance use among jail populations. This integrated view is essential for designing holistic interventions that address the root causes of justice involvement, especially among vulnerable groups like those experiencing homelessness or behavioral health crises.

Ultimately, the JBRS is a foundational component of Washington's broader efforts to modernize and unify criminal justice data. It enables a more transparent, data-informed approach to managing local jail populations and understanding the dynamics of pretrial detention. With improved participation, data quality, and system integration — supported by initiatives like the JDW — JBRS can serve as a powerful tool for advancing equity, reducing jail overcrowding, and promoting fairer justice practices throughout the state.

Caseload Forecast Council

The Caseload Forecast Council's (CFC) Judgment and Sentence (J&S) data is a crucial dataset that captures detailed information about felony convictions in Washington state. This data is derived from standardized judgment and sentencing documents submitted by Washington's Superior Courts following a felony conviction. The J&S data includes comprehensive information such as the type and severity of the offense, criminal history scores, sentence lengths, sentencing enhancements, and whether sentences are served concurrently or consecutively. These records provide a detailed view of sentencing decisions, making the dataset a foundational component for understanding how Washington's sentencing laws are applied in practice.

One of the key strengths of the J&S dataset is its role in forecasting prison and supervision populations. The CFC uses this data to estimate the number of individuals expected to enter or remain within the state prison system or under community supervision in future years. This forecasting supports budget development and resource planning for the Washington State Department of Corrections and other justice system agencies. Because J&S records reflect actual sentencing decisions rather than just charges or bookings, they offer a more accurate basis for understanding long-term incarceration trends and future correctional needs.

The J&S data is also essential for evaluating the consistency and equity of sentencing practices across the state. Researchers and policymakers can analyze the dataset to identify disparities in sentencing outcomes based on race, ethnicity, geographic location, or type of offense. It can reveal, for instance, whether similar cases result in different sentences in different counties or among different demographic groups. This type of analysis can inform policy reforms aimed at promoting sentencing equity and reducing unwarranted disparities, aligning with broader justice reform goals.

Despite its value, the J&S dataset has some limitations related to completeness and linkage. Not all sentencing documents are submitted in a timely or standardized fashion, and some may lack key details necessary for robust analysis. Additionally, while the dataset provides a static view of sentencing at the time of conviction, it does not track subsequent modifications, such as sentence reductions, appeals, or changes to supervision status. Moreover, linking J&S data to other datasets — such as arrest records, court filings, or incarceration data — is often challenging due to a lack of common identifiers, limiting the ability to follow individuals across the justice continuum.

As Washington moves toward more integrated and transparent justice data systems — such as through the JDW — the J&S dataset plays a pivotal role. By incorporating J&S records into broader data integration efforts, interested parties can better understand how sentencing decisions relate to earlier stages of the justice process (like charging or plea bargaining) and to outcomes post-sentencing (like reentry success or recidivism). This integration enhances the ability to conduct longitudinal, equity-focused research and supports the development of evidence-based sentencing policies that are fair, efficient, and aligned with public safety goals.

Washington State Administrative Office of the Courts

The Washington State Administrative Office of the Courts (AOC) serves as the central repository for data collected from courts across the state, including superior, district, municipal, and juvenile courts. The AOC manages and maintains comprehensive records related to court case filings, dispositions, sentencing, bail and bond decisions, warrants, and protection orders. These datasets are essential for understanding how cases are processed through the legal system and for analyzing outcomes at various decision points. The data also includes information on the parties involved — such as charges filed, representation status, and case types — which helps shed light on broader justice system trends.

AOC data plays a critical role in transparency, accountability, and policy development within the judicial system. Policymakers, judicial leaders, researchers, and advocates use this information to examine case processing times, identify disparities in sentencing or bail decisions, and evaluate the impact of new laws and judicial procedures. For example, researchers might analyze how pretrial detention decisions differ based on race or offense type, or whether diversion programs are equitably offered across jurisdictions. This data also helps the courts themselves improve operations by identifying inefficiencies or backlogs in case processing. One of the primary uses of AOC data is to inform legislative and policy decisions. By analyzing the data on case filings, resolutions, and court processing times, policymakers can identify patterns and areas where the justice system may need reform or additional resources. For

example, the AOC tracks the volume of cases in various categories, helping lawmakers gauge where interventions such as funding increases or policy adjustments are needed. Furthermore, data on case outcomes can guide decisions related to sentencing guidelines, bail reform, and other justice-related initiatives. The AOC's ability to provide detailed insights into these areas helps ensure that Washington's justice system remains efficient and responsive to the public's needs.

However, significant challenges exist in terms of data standardization and completeness across courts and jurisdictions. Because Washington's courts operate independently under the state's decentralized judicial system, data quality and reporting practices can vary widely. Not all courts use the same case management systems, and some may enter data inconsistently or omit critical fields. Additionally, many municipal courts are not mandated to report data to the AOC, creating gaps in statewide judicial datasets. These issues hinder the ability to perform accurate statewide analyses and obscure important patterns, especially at the local level.

Another limitation is that AOC data often exist in silos, separate from other criminal justice data systems such as law enforcement or corrections. Without integration, it is difficult to track an individual's journey through the justice system or assess the broader impacts of judicial decisions. For example, analyzing how court decisions affect incarceration rates, recidivism, or access to treatment programs requires linking court data to Department of Corrections or behavioral health datasets. Improved data sharing and integration through platforms like the JDW would significantly enhance the utility and value of AOC data for cross-sector policy analysis.

Despite these challenges, the AOC's data remain one of the most vital components of Washington's justice system infrastructure. Efforts are underway to modernize data systems, standardize reporting practices, and promote greater transparency and accessibility. With robust, integrated, and reliable court data, the state can better identify disparities, evaluate justice reforms, and ensure that legal processes are fair, efficient, and equitable for all Washington residents.

Washington State Department of Corrections

The Washington State Department of Corrections (DOC) maintains extensive datasets on individuals under its supervision, including those incarcerated in state prisons and those on community supervision (formerly known as parole or probation). This data includes demographic information, conviction details, sentencing length, facility location, custody level, program participation, disciplinary infractions, and release information. For individuals on community supervision, data may also include compliance with supervision conditions, drug testing results, and violation reports. The DOC's datasets are essential for tracking trends in incarceration, recidivism, and rehabilitative program outcomes.

DOC data supports a wide array of internal and external uses, including prison operations, policy evaluation, legislative impact assessments, and criminal justice research. Internally, it helps the department manage capacity, assess safety risks, allocate resources, and plan rehabilitative programming. Externally, DOC data is frequently used by researchers, policymakers, and oversight bodies to examine issues such as sentence lengths, population demographics, the effectiveness of reentry programs, and disparities in incarceration by race or geography. This data also informs efforts to reduce prison populations and improve public safety through alternatives to incarceration.

However, challenges persist in accessing and integrating DOC data across broader justice system datasets. For example, linking DOC data with court or arrest records is often complicated by the lack of standardized identifiers or interoperable data systems. Moreover, DOC datasets may not always be structured to support cross-agency analysis or be easily shared due to privacy concerns or outdated

infrastructure. These barriers limit the ability of researchers and policymakers to gain a complete picture of an individual's journey through the justice system — from arrest to incarceration to reentry.

Improving the integration of DOC data with other systems — such as court data, behavioral health services, and community-based programs — would enhance the state's ability to perform longitudinal and equity-focused analyses. For example, understanding how individuals with behavioral health needs fare during and after incarceration could inform diversion efforts and better support reentry. It would also help in identifying patterns related to recidivism, program effectiveness, and the long-term outcomes of correctional interventions. This kind of analysis is especially crucial for ensuring that interventions are equitable and that marginalized populations are not disproportionately impacted.

As Washington moves toward a more integrated criminal justice data infrastructure, such as the JDW, incorporating DOC data is essential. Doing so will allow interested parties to trace individual-level interactions across justice touchpoints and better understand systemic inefficiencies or gaps in services. By enhancing the completeness, accessibility, and interoperability of DOC data, the state can promote more effective, data-informed policy decisions that improve public safety, reduce recidivism, and support successful reintegration into communities.

Discussion

The JDW marks a transformative shift in how Washington state approaches criminal justice system reform. Historically, like many states, Washington's criminal justice and legal systems have operated in fragmented silos — federal, state, local, and Tribal agencies each maintaining separate and often incompatible data systems. This disjointed structure has limited the state's capacity to assess system-wide outcomes, monitor disparities, or implement coordinated interventions across jurisdictions. The JDW represents a vital step forward in addressing these systemic inefficiencies and promoting a unified, data-informed approach to justice.

By centralizing and integrating data from across the justice continuum — including arrest, adjudication, incarceration, and community supervision — the JDW enables longitudinal analysis of individual trajectories through the system. This capability is critical for identifying systemic bottlenecks, and inequitable treatment patterns. Importantly, the JDW fosters transparency and accountability, providing interested parties — from policymakers to the public — with access to standardized, meaningful data that was previously inaccessible or incomplete. This is particularly valuable for understanding and addressing racial, geographic, and socioeconomic disparities in justice outcomes.

The JDW's potential extends beyond the justice system itself. As a cross-sector repository, it promotes collaboration between criminal justice, behavioral health, social services, housing, and public safety agencies. This holistic approach is essential, as many justice-involved individuals experience overlapping challenges related to poverty, mental illness, substance use, and housing instability. Integrating data across these domains supports the development of human-centered interventions that address root causes of system involvement, rather than merely responding to symptoms. For example, early identification of high-risk individuals can inform diversion programs, pretrial support initiatives, and reentry planning, improving outcomes while reducing recidivism.

From a research and policy standpoint, the JDW enables rigorous and innovative inquiry into justice system dynamics. It provides the foundation for applied research that can inform real-time decision-making, monitor the impact of legislative reforms, and support evidence-based program evaluation. This

level of analysis is essential for crafting policies that are responsive to the lived experiences of impacted communities and effective in advancing equity and fairness. Furthermore, the JDW empowers Tribal and local jurisdictions — often under-resourced and excluded from state-level data initiatives — to participate meaningfully in statewide justice reform efforts.

Technological integration through the JDW also supports the development of tools such as real-time data dashboards, case management systems, and coordinated care networks that improve service delivery and agency responsiveness. These advancements are crucial in a system in which decisions at one stage — such as pretrial detention or sentencing — can have profound ripple effects throughout an individual's justice journey.

Cross-sector analyses facilitated by the JDW illuminate how broader social determinants — such as education, employment, and health — interact with justice involvement. This understanding is essential for creating comprehensive, community-based strategies that prevent system entry, reduce reliance on incarceration, and promote restorative justice. By demonstrating the fiscal and social benefits of upstream investments and community supports, the JDW also strengthens the case for sustainable, equity-driven reforms.

In summary, the JDW represents a paradigm shift toward an integrated, data-driven justice system in Washington. It breaks down traditional silos, enabling a holistic view of system performance and individual experiences. This integrated approach enhances transparency, supports cross-sector collaboration, and equips interested parties with the tools needed to design and implement reforms that are just, effective, and accountable. Through its continued development and use, the JDW has the potential to fundamentally reshape the state’s approach to criminal justice — centered on equity, informed by data, and committed to meaningful change.

Appendix A: JDW Overview

Table 1 shows the data tables that are integrated into the JDW by agency, table name and purpose, years covered, and number of data elements. Table 2 dives deeper by showing the data elements of the JDW by agency, table name, variables name, definition, and data type.

Table 1. JDW Data Table

Agency	Table/File Name	Table/File Purpose	Years Covered	Data Variables	Table Granularity
Administrative Office of the Courts (AOC)	Bulk Case and Charge Data	includes case level data (court name and level; case num, type, and cause; charge date, law and result; sentence information)	01/01/2010-present	26	all cases seen by courts in WA
Administrative Office of the Courts (AOC)	Bulk Case and Person Data	includes demographic details (name, alias, race/ethnicity, DOB, etc.) and the case (court name and level, case num, type, etc.) associated to them	01/01/2010-present	16	all cases seen by courts in WA
Caseload Forecast Council (CFC)	Sentence	includes data about information associated with the sentence imposed; each record contains information from a J&S.	01/01/2000-present	34	Individual-level sentenced-based data (**J&S**) from each county for every felony sentence imposed in the state
Caseload Forecast Council (CFC)	Current Offenses	includes data about current offense; each record represents the individual count level within one sentencing document and contains the sentencing on the current offenses.	01/01/2000-present	13	Individual-level sentenced-based data (**J&S**) from each county for every felony sentence imposed in the state
Caseload Forecast Council (CFC)	Other Current Offenses	includes data about other the current offense(s); each record represents an offense at the count level for convictions entered or sentenced on the same date as the conviction(s)	01/01/2000-present	5	Individual-level sentenced-based data (**J&S**) from each county for every felony sentence imposed in the state
Caseload Forecast Council (CFC)	Offender	Includes the demographic information of the person for each sentence imposed in the fiscal year, as reported on J&S	01/01/2000-present	7	Individual-level sentenced-based data (J&S) from each county for every felony sentence imposed in the state
Caseload Forecast Council (CFC)	Historical Offenses	includes data about criminal history; each record within the table represents an offense in the individual's criminal history (convictions existing before the date of the current conviction).	01/01/2000-present	5	Individual-level sentenced-based data (**J&S**) from each county for every felony sentence imposed in the state
Caseload Forecast Council (CFC)	Exceptional Sentence Reasons	includes data about exceptional sentence reasons; the court may impose a sentence outside the	01/01/2000-present	2	Individual-level sentenced-based data (J&S) from each

		standard sentence range for an offense if it finds that there are substantial reasons justifying an exceptional sentence			county for every felony sentence imposed in the state
Department of Corrections (DOC)	Person_xxxx	includes data about the DOC individual - demographic details such as name, alias, race/ ethnicity, homelessness, etc.	04/01/1953-present	66	Incarcerated Individuals in WADOC jurisdiction
Department of Corrections (DOC)	Programs_xxxx	includes data about the programs (e.g., SOTAP, PIO) that the DOC individual participate(d/s) in	04/01/1953-present	19	Incarcerated Individuals in WADOC jurisdiction
Department of Corrections (DOC)	AdmitReleases_ xxxx	includes data about a DOC individual's admissions, releases, and movements	04/01/1953-present	45	Incarcerated Individuals in WADOC jurisdiction
Department of Corrections (DOC)	Crime_xxxx	includes data about a DOC individual's crime/offense related to DOC custody	04/01/1953-present	22	Incarcerated Individuals in WADOC jurisdiction
Department of Corrections (DOC)	Custody_xxxx	includes data about a DOC individual's custody level information	04/01/1953-present	11	Incarcerated Individuals in WADOC jurisdiction
Washington Association of Sheriffs & Police Chiefs (WASPC)	WASPC_RECID_ REPORT_20xx-xx-xx	provides individual level booking and release records from most city and county jails in Washington	09/01/2009-present	35	jail and booking releases on an individual level basis