

---

**Month 202X**



# **Electronic Data Extract Guide**

## **Jails and Pretrial Services Agencies**

Prepared by

**Matthew DeMichele**

**Suzanne Strong**

**Milton Cahoon**

RTI International

3040 E. Cornwallis Road

Research Triangle Park, NC 27709

OMB Control Number: ##### Expiration Date: #####

---

# Contents

---

<b>Section</b>	<b>Page</b>
<b>1. National Pretrial Reporting Program Overview</b>	<b>1-1</b>
<b>2. Data Preparation Instructions</b>	<b>2-1</b>
2.1 Identifying Eligible Cases.....	2-1
2.2 File Structure.....	2-2
2.3 File Format.....	2-2
2.4 Supporting Documentation.....	2-2
<b>3. Data Submission Instructions</b>	<b>3-3</b>
3.1 Will the data be secure and kept confidential?.....	3-4
3.1.1 Data Transmission.....	3-4
3.1.2 Data Storage and Access.....	3-4
3.1.3 Data Publication.....	3-5
3.2 When is the submission due?.....	3-5
3.3 What if I am unable to provide all the requested data?.....	3-5
3.4 What happens after we submit the data?.....	3-5
3.5 Who do I contact if I have questions?.....	3-6
<b>Appendix</b>	
<b>Appendix A: Requested Data Elements, Definitions, and Formats</b>	<b>1</b>

---

## 1. National Pretrial Reporting Program Overview

The goal of the Bureau of Justice Statistics' (BJS) National Pretrial Reporting Program (NPRP) is to collect information on persons charged with felony cases in state courts, and to collect contextual data on those persons from case filing to case disposition and sentencing. The NPRP will collect case-level information on pretrial release and detention, financial and other conditions associated with pretrial release, and any failures to appear, technical violations, or new arrests that occur during pretrial release.

The project will be completed in several phases by RTI International, with support from project partners National Center for State Courts (NCSC), National Association of Pretrial Services Agencies (NAPSA), Applied Research Services (ARS), and Pragmatica, Inc. The first phase collected information from states and counties about the systems that may collect and store these data. Such systems included courts, local jails, and pretrial services agencies. Data capacity surveys were conducted with each agency in the largest 200 counties in the U.S. The surveys allowed RTI and BJS to identify the data providers, understand their systems capabilities, and identify variables that potentially could be used to link the data across agencies.

The remaining phases involve a staged data collection from the 75 largest counties, and a sample of 50 of the remaining 125 counties. RTI and BJS have developed a sampling plan to maximize responsiveness and data representativeness and minimize respondent burden and project costs.

## 2. Data Preparation Instructions

This section outlines how to prepare and submit your NPRP data extract submission.

### 2.1 Identifying Eligible Cases

**Jails.** Please include one record for each admission of an offender arrested for and booked into the jail for at least one felony charge in calendar year 2019. If the offender appears in a later admission with a felony charge, please include that as a separate admission. If your jail management system organizes bookings in some other way, please contact the NPRP staff listed in Section 3.5.

**Pretrial services agencies.** Please include one record for each defendant with at least one new felony charge filed in your system as a new case in calendar year 2019, even if that person was already a client. If you are unable to determine whether the charges included a felony at the opening of the case, please contact the NPRP staff listed in Section 3.5.

---

## 2.2 File Structure

BJ5 is interested in a person-case and following that person-case from arrest and booking into jail, through pretrial processing to court case outcomes. This is an individual- and case-level data collection. You can provide data in any format that is convenient for you, but we ask that you please provide supporting documentation, if available.

## 2.3 File Format

**There is no required format for the data you submit; use whatever is most convenient for you.** All file formats will be accepted. Some common file formats include:

- Text files (fixed width, delimited)
- Excel or .csv files
- Access database extracts
- SQL server database
- Data analysis software files (e.g., SAS, STATA, SPSS, or R data files)

The suggested coding classifications and value labels are provided in **Appendix A**. You are not asked to recode or manipulate your data prior to submission. If you have questions about any of the data elements requested, please contact the persons listed in Section 3.5. We recognize that systems vary in terms of the ability to store, extract, and share data, and we are prepared to assist you.

**Appendix A** is a guide as to how we expect to recode and standardize the data you submit so that jail, pretrial services, and court data all have similar formats. Additionally, data from your jurisdiction will be combined with data from up to 75 other counties to get a robust perspective of pretrial case processing in the U.S. During the data standardization process, RTI and NCSC may be in contact with jurisdiction personnel to request clarification on data fields and meanings to ensure that all submitted data are processed correctly.

## 2.4 Supporting Documentation

If possible, we ask that you provide supporting documentation with your submission. Specifically, we request:

- Date range of the data (e.g., 01/01/2019 through 12/31/2019)
- Date that the data extract was pulled
- Data point of contact (POC) (i.e., name, organization, address, telephone, and email address)
- File format of the data extract
- Known data limitations or quality issues
  - Missing data:

- 
- System-missing (requested data element is not available in the system)
  - Unit-missing (requested data element is available, but mostly blank or missing)
  - o Other common data issues include
    - Misspellings
    - Redundancy or duplication (e.g., two date fields for one event)
  - Data formatting information
    - o Data dictionaries, including variable/column names, variable description, expected variable values
    - o Any known discrepancies in the names of data elements in **Appendix A** and how your system labels the data elements

### 3. Data Submission Instructions

RTI will create a private, password-protected user account for each Data POC to upload data to Amazon Web Services (AWS) Simple Storage Service (S3). This AWS S3 storage location will only be accessible from designated network subnets. The Data POC will need to provide their subnet range or specific IP address from which they will be accessing AWS S3 from so the access control rules may be updated to grant access to the Data POC from their network. A free and easy way to discover your IP address is to go to <https://www.iplocation.net/> Once data access has been set up, the Data POC will log into their private account to transfer the requested data to a secure central data storage system on AWS S3.

Data security note: All data are encrypted in transit to AWS and at rest within AWS (SSL in transit and AES 256-encryption at rest), complying with the FIPS 140-2 standard. The secure AWS S3 repository will hold all raw data files received from the courts, jails, and pretrial services agencies until they are processed, linked, de-identified, and subsequently deleted by RTI data analysts. RTI controls access to the data storage system; all access to data resources will be logged, and the entire infrastructure will be reviewed and regularly scanned for vulnerabilities. The data storage system will be configured to deny public access by default, and we will use Amazon's Macie service to regularly scan and evaluate the security status of the storage. All RTI and NCSC staff granted access to data files (identified and deidentified files) will be required to sign a Staff Data Security Agreement. This pledge outlines staff responsibilities for protecting the confidentiality of all information identifiable to a private person that is collected during the project. The RTI Principal Investigator is responsible for maintaining up-to-date record of signed pledges.

---

### **3.1 Will the data be secure and kept confidential?**

Consistent with its statutory obligations (34 U.S.C. § 10134), BJS only uses information collected under its authority for statistical or research purposes. Further, BJS is required by law to protect the confidentiality of all personally identifiable information (PII) it collects or acquires in conjunction with BJS-funded projects (34 U.S.C. § 10231), and must maintain the appropriate administrative, physical, and technical safeguards to protect the identifiable information against improper use or unauthorized disclosure. BJS will not use or reveal data identifiable to a private person, except as authorized under 28 CFR § 22.21 and § 22.22. The BJS Data Protection Guidelines summarize the federal laws, regulations, and other authorities that govern information acquired under BJS's authority, and are published on the BJS website: [https://www.bjs.gov/content/pub/pdf/BJS\\_Data\\_Protection\\_Guidelines.pdf](https://www.bjs.gov/content/pub/pdf/BJS_Data_Protection_Guidelines.pdf). RTI and NCSC are required to adhere to these same requirements as a condition of funding.

#### **3.1.1 Data Transmission**

RTI and NCSC project staff will receive data sets in a secure manner via an encrypted AWS GovCloud S3 server, appropriate for files with PII. RTI will create a private, password-protected user account that relies on an email address and IP address for each agency to upload data to the AWS S3 storage location. This AWS S3 storage location will only be accessible for the email and IP address granted access to the server. Any data set(s) electronically transmitted to BJS will be over the DOJ's Office of Justice Programs (OJP) secure transfer site.

#### **3.1.2 Data Storage and Access**

The secure AWS S3 repository will hold all raw data files received from the agencies until they are processed, linked, and subsequently deleted by RTI. Once received and linked with other available records, raw data files will be stripped of PII and replaced with an anonymous identifier. RTI will retain a PII-anonymous identifier crosswalk in case a revised file is submitted later. This file will be encrypted, password protected, and stored on a secure RTI server in the event the AWS S3 storage location is breached. Access to the S3 instance will be restricted to individuals with an identified business need. RTI controls access to the data storage system; all access to data resources will be logged, and the entire infrastructure will be reviewed and regularly scanned for vulnerabilities. PII is encrypted while in transit, and access to the data will be limited to those employees who have a need for such data and have signed a confidentiality pledge. The pledge includes an agreement to comply with all data security and human subjects' protection requirements.

---

### **3.1.3 Data Publication**

BJS only publishes de-identified data at the aggregate level in its project findings, reports, data files, and other statistical products. BJS archives its published data and related data documentation (e.g., user guides) at the National Archive of Criminal Justice Data (NACJD), located at the University of Michigan. To the extent practical, BJS removes, masks, or collapses direct and indirect identifiers prior to sending data to NACJD to protect confidentiality. NACJD takes additional precautions to protect confidentiality, including conducting a comprehensive disclosure risk review to determine the appropriate level of security that should be applied to the data. For more information on data requiring additional security protections, please see:

<https://www.icpsr.umich.edu/icpsrweb/content/NACJD/restricted.html>.

We understand that your agency may have preexisting policies in place around data sharing, and we will work with your agency to meet any data transfer or agreement requirements you may have. While each of the identified data elements was selected to help fully understand the processing of case data, we recognize that not all data elements may be collected or readily available electronically for public use. Please notify us if you limit the amount or type of data you can release.

### **3.2 When is the submission due?**

We ask that all participating agencies provide their data by ##/##/####. However, please reach out to the staff identified in Section 3.5 if you need additional time to submit your data.

### **3.3 What if I am unable to provide all the requested data?**

The data extract guide is intended for use by pretrial services agencies and jails. We do not expect that any one organization has all of the data elements requested in Appendix A. If your agency does not collect any of the information in Appendix A, or it would be too burdensome to provide the data, please contact the staff in section 3.5.

### **3.4 What happens after we submit the data?**

RTI or NCSC will review the contents of the data files and conduct a series of checks to the data elements requested in the **Appendix A**. This should be completed within 2-4 weeks of submission. RTI or NCSC will then contact the Data POC to review and confirm the findings from the review. We may also have questions about variable values or labels and will take all steps to understand your data submission.

---

### **3.5 Who do I contact if I have questions?**

You may reach out to any of the following RTI staff members for questions or support in submitting your data:

- Marianne Ayers – NPRP Data Acquisition Support
  - o Email: [mayers@rti.org](mailto:mayers@rti.org)
  - o Phone: 1-800-334-8571 (ext 23841)
- Milton Cahoon – NPRP Data Acquisition Lead
  - o Email: [mcahoon@rti.org](mailto:mcahoon@rti.org)
  - o Phone: 1-800-334-8571 (ext 27161)
- Suzanne Strong – NPRP Project Director
  - o Email: [ssstrong@rti.org](mailto:ssstrong@rti.org)
  - o Phone: 1-800-334-8571 (ext ### ##)



## Appendix A: Requested Data Elements, Definitions, and Standard Formats

The following series of tables includes the data elements for the NPRP. The name of the variable and the description provided should help you to locate a similar data element in your data management systems. We are also including a standard format, which is how we expect to standardize the data you submit.

Please contact anyone in section 3.5 if you have any questions about the following data elements.

**Table 1. Individual Identifiers and Demographic Data**

Name	Definition	Standard Formats
First name	The individual's first name	Text, character, string
Middle name (if available)	The individual's middle name	Text, character, string
Last name	The individual's last name	Text, character, string
State ID number	The individual's unique, fingerprint-supported state identification number	
FBI Number	The unique identification number given by the Federal Bureau of Investigation's Interstate Identification Index to each offender (if available)	
Other personal identifiers if SID and FBI number are not available	If SID and FBI number are not available, please provide any other unique identifier assigned to a person, such as a Social Security Number or Driver's License Number	
Other system identifiers assigned to a <b>person</b>	Include any other unique identifiers that specify the individual in your data system. For example, booking ID, inmate ID, or client ID	

Name	Definition	Standard Formats
Date of birth or age	The individual's date of birth. If not available, individual's age (and please include age at what reference period in the documentation. For example, age at arrest, age at booking, age at filing)	mm/dd/yyyy
Sex	The individual's biological sex/sex assigned at birth	Male, Female, Other, Unknown
Race	The individual's race(s)	(OMB race categories) White, Black or African American, Asian, Native Hawaiian or Other Pacific Islander, American Indian or Alaska Native, Other, Unknown.
Ethnicity	The individual's ethnicity	(OMB ethnicity categories) Hispanic or Latino; Not Hispanic or Latino; Unknown

**Table 2. Jail/Detention Information**

<b>Name</b>	<b>Definition</b>	<b>Standard Formats</b>
Date of offense	Date the alleged offense occurred	mm/dd/yyyy
Date of arrest	Date the individual was arrested for the offense	mm/dd/yyyy
Arrest charge(s)	Arrest charges or the charges for which the person is incarcerated pretrial	State statute number, text description
Arrest charge level	The level of charges. For inclusion in NPRP, the individual should have at least one felony charge	Felony, Felony A, Misdemeanor, Misdemeanor II
Date of booking into facility for arrest charges	Date individual was booked into the jail for the arrest for the current case or charge	mm/dd/yyyy
Date of pretrial risk assessment(s)	Date the pretrial risk assessment conducted. If multiple assessments are conducted for one individual, please include all, with dates. Include description in data documentation	mm/dd/yyyy
Risk assessment score	The risk assessment category assigned to the defendant. If multiple scores are available for one individual, please include all. Please include a description of the risk assessment score in the data documentation	
Date bond set by court or other official	The date the bond was set by court, law enforcement, or other justice official	mm/dd/yyyy

Name	Definition	Standard Formats
Type of bond set	The type of bond specified by the court, law enforcement, or other justice official. Please include in the data documentation the types of bonds available and how they are recorded in your system	Cash, percentage, surety, property, personal recognizance, unspecified
Type of detention ordered	Court ordered no bond or no release. We would like to know if person is held because court ordered no bond as compared to a person held because a bond is ordered but not posted	No bond
Amount of bond	If a financial bond is ordered, include the bond amount	Numeric
Conditions of bond	Any conditions ordered for the bond. Please include in the data documentation the available bond conditions that are tracked in your system	No contact with victim, drug and alcohol testing, drug and alcohol treatment, curfew, maintain employment, other, etc
Date bond posted	If possible, the date the bond was posted	mm/dd/yyyy
Type of bond posted	If possible, the type of bond posted	Surety/bail bonds company, cash bond, property bond, other
Date of pretrial release from facility	Date individual was released from facility (pretrial). Please note in documentation if this date can be determined, or if there is only one field for any type of release	mm/dd/yyyy

Name	Definition	Standard Formats
Pretrial release reason	If offender was released pretrial, how offender was released	Posted bond, released on recognizance, transferred to house arrest/electronic monitoring, released to pretrial supervision
Date of final release from jail, or date defendant changed from pretrial status to sentenced/convicted status	The date the jail released or changed the status of the defendant after a final verdict was entered in the court case (e.g., conviction, dismissal)	mm/dd/yyyy

**Table 3. Pretrial Supervision, Failures to Appear, New Arrests, and Technical Violations**

Name	Definition	Standard Formats
Date agency began supervising client	The date the agency began supervision of the defendant	mm/dd/yyyy
Charges	If available, the defendant’s charges. Please include in the documentation the source of the charges (e.g., arrest charges, court filing)	State statute number, text description
Charge level	If available, the level of charge. For inclusion in NPRP, the defendant should have at least one felony charge	Felony, Felony A, Misdemeanor, Misdemeanor II
Date of pretrial risk assessment(s)	Date the pretrial risk assessment conducted. If multiple assessments are conducted for one defendant, please include all, with dates. Include description in data documentation	mm/dd/yyyy
Risk assessment score	The risk assessment category assigned to the defendant. Please include a description of the risk assessment score in the data documentation	
Level of pretrial supervision	The level of supervision determined by the court or the supervising agency. These vary greatly; please include in the data documentation the method used to determine how intensely you monitor clients	

Name	Definition	Standard Formats
Conditions of supervision	Include all conditions of supervision ordered by the court. If possible, please list additional conditions added by your agency separately from those ordered by the court	In-person reporting, telephone reporting, home visits, curfew, other
Violation(s) of supervision	How the defendant violated supervision, if applicable. There should be one type of violation entered per occurrence. If a defendant incurs multiple violations on the same date, list each violation separately. Include the violations tracked by your agency in the data documentation	Fail to report, fail to comply, fail drug test, fail to appear, new arrest, other
Date(s) of violation of supervision	The date(s) the client violated supervision. If the date is for a failure to appear, please indicate the type of hearing missed, if possible	mm/dd/yyyy
Date supervising agent filed a violation report	Date the supervising agent reported the violation to the court or other administrative agency. Please provide data documentation about when and how these reports are filed	mm/dd/yyyy
Outcome of violation of supervision	Outcome of the violation	Continued on supervision, bail revoked
Date of violation of supervision outcome	Date pretrial release was revoked	mm/dd/yyyy
Date pretrial supervision ended	The date the agency stopped supervising the defendant	mm/dd/yyyy

Name	Definition	Standard Formats
Final outcome of supervision	Outcome of the pretrial supervision	Successful case closed, unsuccessful case closed, other

**Table 4. Potential linking identifiers**

Name	Definition	Standard Formats
Court case number	The court case number for the defendant	
Jail person identifier	The inmate id number for the jail	
Pretrial agency case number	The number assigned to the person or to the case under supervision	

**Table 5. Manner of Disposition and Outcomes**

Name	Definition	Standard Formats
Manner of disposition	Type of hearing for the disposition	Plea, court trial, jury trial, other
Type of disposition	Type of disposition for the case or charge	Nolle prosequi, dismissal, acquittal, not guilty, probation before judgment, guilty, Alford plea, no contest plea, other



**Table 6. Sentencing**

<b>Name</b>	<b>Definition</b>	<b>Standard Formats</b>
Total sentence to incarceration – type of facility	Type of facility ordered for the sentence	Prison, jail, other
Total length of incarceration	Length of the sentence to a facility. Please specify the unit of time for the sentence in the data documentation	Number (specify)
Sentence suspended	Whether any time of the sentence ordered is suspended. Please include detail in the data documentation	Y/N
Length of sentence suspended	Portion of the sentence that is suspended	Number (specify)
Credit for time served	Whether the defendant received credit for time served incarcerated pretrial. Please include in the documentation whether electronic monitoring or house arrest counts towards credit time	Y/N
Length of credit for time served	Length of time the defendant received credit for time served. Please include in the data documentation the unit of time	Number (specify)
Total sentence to probation	Indicator whether the defendant was sentenced to probation	Y/N

Name	Definition	Standard Formats
Total length of sentence to probation	Total length of time the defendant was sentenced to probation. Please specify the unit of time for the sentence in the data documentation	Number (specify)
Conditions of probation	Conditions imposed as part of probation	Counseling, drug treatment, drug testing, domestic violence program, anger management, other
Other sentence imposed	Indicator whether another sentence was imposed. Please include details about available sentence types in your data documentation (e.g., community service may be a type of sentence or part of a probation order)	Y/N
Length of other sentence imposed	Length of the other sentence ordered. Please specify the unit of time for the sentence in the data documentation	Number (specify)
Monetary sentence ordered	Court ordered fines as part of sentence. Please include detail in the data documentation how monetary fines are tracked	
Amount of monetary sentence ordered	Amount of the monetary sentence ordered	Number
Restitution	Whether restitution was ordered.	Y/N

---

Name	Definition	Standard Formats
Restitution	If available, was restitution to the victim or the state. Please include detail in the data documentation.	Victim, State
Court costs	Whether court costs were ordered to be paid by the defendant	Y/N