National Archive of Criminal Justice Data: Archiving Orientation for OJJDP Research Grantees

May 22, 2018

Barbara Tatem Kelley, OJJDP
Social Science Analyst and NACJD Program Manager

Chelsea Samples-Steele, ICPSR
NACJD Project Manager for User Experience

Justin Noble, ICPSR
NACJD Project Manager for Data Services
Agenda:

• Guidance for OJJDP Researchers
• Privacy and Human Subjects
• Archiving Requirements
• Overview of Materials to Deposit
• Preparing Data for Sharing
• Data Deposit Demonstration
• What Happens After the Deposit?

NACJD Deposit Review, Curation, and Data Release
Guidance for OJJDP Research Grantees is now available:

https://www.ojjdp.gov/research/research-grantee.html

- Privacy and Human Subjects Protections
- Research Participant Support Costs and Incentives
- Research Independence and Integrity
- Press Releases and Publications
- Final Reports for Research Grantees
- Copyright and Data Rights
- Data Archiving
- Key Personnel
- Progress Reporting
- Research Award Closeout
- Budget and Other General Grant Requirements and Resources
Researchers must address **privacy** and **human subjects protection** over the entire **research lifecycle** – from planning of the study through data archiving.
Privacy Certificate:

- Department of Justice (DOJ) regulations (28 C.F.R. Part 22) require recipients of OJP funding to submit a Privacy Certificate as a condition of approval of any grant application or contract proposal that contains a research or statistical component under which "information identifiable to a private person" will be collected, analyzed, used, or disclosed.

- Include explicit statement that de-identified data will be archived at NACJD.
Human Subjects Protection:

- To protect human subjects of federally funded research, DOJ regulations (28 C.F.R. Part 46) require that an Institutional Review Board (IRB), in accordance with the regulations, review and approve most research involving human subjects that any federal department or agency conducts or supports before an award recipient may expend federal funds for that research.
Data Release:

- The research grantee will make no guarantee, without prior OJJDP approval, that the data collected, acquired or produced as part of this project will not be transferred or released.

- Such a guarantee would preclude OJJDP from archiving and making available OJJDP-funded data.
Data Archiving Plan and Purpose:

- A data archiving plan will be due to OJJDP within 6 months of the grant period start date. It will include a specific schedule, tasks, and milestones. Please upload as a special report in GMS.

- Data are not only archived for safekeeping but also made accessible for secondary analysis.
Data Archiving Requirement:

• The grantee agrees to deliver a public-use dataset containing all data collected under this project to the National Archive of Criminal Justice Data (NACJD) unless notified in writing by OJJDP that they are exempt from this requirement.

• As of Fiscal Year 2016, all research grantees are to complete this requirement 30 days before the end of the project period.
Overview of Materials to Deposit

I. Data (de-identified)

II. Technical Documentation
   – Codebooks
   – Syntax or special coding
   – Data collection instruments

III. Study Documentation
    – IRB approval
    – Informed Consent
    – OJJDP Privacy Certificate
    – OJJDP Data Archiving Plan
    – Final Report

IV. Bibliographies
I. Data File(s)

• Make sure personally identifying information is removed (e.g., name, address, phone number, etc.)

• Include raw data AND any transformed variables, such as computed scales or recoded variables (age groups, family type, etc.)

• Preferred formats:
  • Quantitative: SPSS (but SAS, Stata, Excel, etc. accepted)
  • Qualitative: MS Word (but other formats accepted)
Files containing **quantitative** data should:

- Have a unique identifier assigned to every record
- Use mutually-exclusive, exhaustive and precisely defined codes
- Distinguish between different types of missing data. For example:
  - -9 = item non-response
  - -99 = inappropriate response/value
  - -999 = valid skip/NA
  - -9999 = did not participate in module/wave
- Include all cases and variables
• Files containing **qualitative** data should:

• Use pseudonyms or other generalized text before data are archived

• Transcribed data should:
  • Contain unique identifiers
  • Follow a uniform layout as well as header, page numbers, line breaks, and speaker tags indicating the question-answer sequence
II. Technical Documentation

• Codebooks for each dataset

• Syntax files and any notes about coding decisions

• Data collection instruments (unused, original questionnaires, protocols, interviewer guides, coding instruments, etc.)

• User guide (recommended)

• Desired formats: PDF, MS Word, Rich Text, Plain text
Codebooks

- Variable names & labels
- Variable value labels
- Missing data designations
- Variable formats
  - Variable type
  - Variable length
  - Decimal specifications
- Exact descriptions of the data elements (wording of questions, items, records, etc.)
- Skip patterns
- Univariate statistics
- Column locations for fixed-length ASCII data files
Syntax Files

- Syntax used to recode variables
- Syntax used to compute scales, constructs, and new variables
- Syntax used to merge files (if applicable)
- Syntax used to create weights
- Coding schemes for qualitative data
- Analysis syntax used to generate descriptive and other statistics in the final report
- Feel free to include any annotation within syntax files
User Guide (Recommended)

Suggested User Guide Elements include:

- An inventory of data and documentation files
- Imputation procedures
- Coder manual to describe naming conventions, skip patterns, missing data, etc.
- Information on how to aggregate, merge, and/or subset the data
- Errata identifying known mistakes
- FAQs
III. Study Documentation

- IRB documentation
- Consent forms
- OJJDP privacy certificate
- OJJDP data archiving plan
- Final report
Data Archiving Plan

- Data description
- Format
- Security
- Responsibility
- Access and sharing
- Ethics and privacy
- Intellectual property
- Identify dataset(s) and variables used
- Metadata
IV. Bibliography of Project Publications

• **Citation information** for all publications generated *using the data* from this project (e.g., journal articles, book chapters, reports, working papers, conference presentations or abstracts, and posters)

• Include the **final report** as one of the publications

*NACJD and ICPSR maintain a bibliography of literature linked to the datasets in their collections. The data-related bibliography assists in the (a) discovery and (b) proper secondary use of the data*
Any questions about requirements?

NACJD staff is ready to assist you with the deposit process

1-800-999-0960

nacjd@icpsr.umich.edu
Before You Start the Deposit Process

- Locate the required files: data, codebook, etc.
- Use a checklist to make sure you have everything
- Make sure you have the most recent or authoritative versions of the files
- Access the online deposit platform from the NACJD (ICPSR) website
- Note that the deposit system allows you to save your work: **You do not have to complete the deposit in one sitting**
Data Deposit Demonstration
What Happens After the Deposit?

1. Initial Review
2. Communication with you and/or your grant manager if needed
3. Full Review/Data Curation
Data Curation Process

• **Goals**
  • Ensuring quality and confidentiality
  • Enhancing data discovery
  • Facilitating secondary analysis

• **Main steps**
  • Deposit and disclosure review
  • Metadata and documentation
  • Product suite: SPSS, Stata, SAS, etc.

• **Dissemination**
  • Your data will be included in a searchable collection of 2,700+ NACJD (10,500+ ICPSR) datasets
A Note on Disclosure Assessment and Protection

**Distribution Policy:** NACJD does not publicly release data containing information that might reasonably identify any individuals, organizations, or agencies who provided information under a promise of confidentiality.

To attain our goal of building data and documentation for greatest access and least risk, we will take steps necessary to prevent disclosure.
Data Dissemination Modes

- Public download
- Restricted
  - Secure download
  - Virtual Data Enclave (VDE)
  - Physical Data Enclave (PDE)
- Online Analysis
  - Public or restricted
Helpful Resources

1. Guide to Social Science Data Preparation and Archiving:  
   www.icpsr.umich.edu/icpsrweb/content/deposit/guide

2. Creating a Data Archiving Plan:  
   www.icpsr.umich.edu/icpsrweb/content/deposit/guide/chapter1.html

3. Guidelines for Depositing Quantitative Data at ICPSR:  

4. Example Study Home Page with User Guide:  
   www.icpsr.umich.edu/icpsrweb/NACJD/studies/34931?q=34931
Thank You!

Questions?

NACJD
1-800-999-0960
nacjd@icpsr.umich.edu