

## **Protecting Confidential Data at the National Archive of Criminal Justice Data**

The National Archive of Criminal Justice Data (NACJD) employs stringent procedures to protect the confidentiality of subjects while distributing data collections for secondary research and other statistical analysis. These procedures include:

- Detailed review of all datasets to assess disclosure risk upon receipt of the data
- Modifying data if necessary to reduce risk and protect respondents' confidentiality
- Requiring all prospective data users to agree to follow Terms of Use (which include Federal laws and regulations) prior to accessing data
- Restricting access to datasets for which modifying the data would substantially limit utility, or where the risk of disclosure remains high even after removal of identifying information (see more information below)
- Training of staff and consulting with data producers in methods of disclosure risk assessment and mitigation

### **Restricted Access**

Restricted Access is used when an entire study, or part of a study, contains data that cannot be reasonably altered to protect subject confidentiality without significantly degrading the value of the research data. While all data are reviewed, and in some cases modified, to address disclosure risk, the level of access restriction employed for a data collection or part of a collection is determined by NACJD's Director and/or Project Manager in cooperation with the data depositor/principal investigator, the sponsor of the data collection, and the sponsor of the data archiving.

Decisions to restrict access to a dataset are made after a full disclosure risk review of the data, during which staff evaluate the following and other characteristics of the data:

1. The source and nature of the data (e.g., public use information vs. self-report; agency vs. individual);
2. The extent to which all direct, personal identifiers have been masked or removed, regardless of the source and nature of the data;
3. The extent to which indirect identifiers, such as addresses, dates or combinations of characteristics, are present in the data, and if applicable have been recoded (e.g., categories are collapsed) or masked to prevent (re)identification of subjects;
4. The sample characteristics and risks associated with reidentification of subjects.

Following the disclosure risk review, if disclosure risk remains, but further modification to the data would reduce analytic utility of the data, the NACJD Director and/or Project Manager will recommend restricted access.

Restricted studies, or parts of restricted studies, are not downloadable directly from the NACJD website. Researchers who want to use restricted data must complete an online application to request access.

NACJD currently offers the following three levels of restricted access: Restricted Download, Virtual Data Enclave, and Physical Data Enclave. The access mechanism chosen for a given dataset depends on the level of disclosure risk and the sensitivity of the data. Each access mechanism has slightly different application and data access requirements.

## **Restricted Download**

Access to restricted download collections is arranged after the prospective data user submits an online application via ICPSR's Data Access Request System (IDARS), complete with a Restricted Data Use Agreement (RDU). The RDU has several requirements of prospective data users and their institutions, such as:

- certification that the NACJD data will be used only for research or statistical analysis
- agreement to protect the confidentiality of subjects
- explanation of why access to the restricted data is needed
- description of how the data will be securely stored and accessed
- data users must have a PhD or other research-appropriate terminal degree with either a faculty appointment or an appointment that is eligible to be a principal investigator at their institution

The RDU must also be signed by a representative of the data user's institution who is authorized to execute a legal agreement on that institution's behalf. The completed RDU must be accompanied by approval or exemption from the Institutional Review Board at the data user's institution, which indicates that a responsible officer at the institution has reviewed the project and finds the project's terms of use for the requested data to be in compliance with their organization's human subjects and confidentiality rules.

The final determination on access to Restricted Download studies rests with NACJD's Director and/or Project Manager. After approval, the data is compressed and encrypted, then sent directly to the data user. The data user must move the files to the secure location as specified within their IDARS application prior to receiving the data decryption password.

From the date of the fully executed RDU, the user is granted access to the data for two-years. After two years, users are required to either securely destroy any copies of the data and return a notarized affidavit affirming data destruction to NACJD, or submit a request for an extension of the RDU.

## **Virtual Data Enclave (VDE)**

The VDE is a virtual machine launched from the data user's own computer that operates on a remote server. Users access data stored on ICPSR's servers via a remote desktop connection and are unable to move or copy files outside of the secure environment. All output files are reviewed by NACJD staff for compliance with confidentiality standards before being removed from the VDE.

To access data in the VDE, data users must submit an online application via ICPSR's VDE application portal, complete with a fully executed RDU. The completed RDU must also be accompanied by approval or exemption from the Institutional Review Board at the data user's institution, which indicates that a responsible officer at the institution has reviewed the project and finds the project's terms of use for the requested data to be in compliance with their organization's human subjects and confidentiality rules. The final determination on access to VDE studies rests with NACJD's Director and/or Project Manager.

## **Physical Data Enclave (PDE)**

For highly sensitive data, ICPSR has a physical data enclave, which requires that approved

users are on site at ICPSR in Ann Arbor, MI to use the data. Prospective data users must apply to use data in the PDE. Application materials can be submitted to [icpsr-help@umich.edu](mailto:icpsr-help@umich.edu). Upon approval, data users are set up to work on a standalone computer that is not connected to a network or any other devices. Data use in the PDE is monitored by ICPSR staff at all times. All written notes, data analysis output and other electronic files are reviewed by NACJD staff for compliance with confidentiality standards before being removed from the PDE.

All PDE applications must include a fully executed RDUa and approval or exemption from the Institutional Review Board at the data user's institution, which indicates that a responsible officer at the institution has reviewed the project and finds the project's terms of use for the requested data to be in compliance with their organization's human subjects and confidentiality rules. The final determination on access to PDE studies rests with NACJD's Director and/or Project Manager.

### **Restricted Survey Documentation and Analysis**

NACJD also distributes a limited collection of some restricted-use data via a password-protected secure online analysis system. This system allows data analysis on the web without the need to download entire datasets. This method does not provide access to the microdata, and limits the kinds of analysis that can be performed and output that is provided. Users interested in accessing restricted-use data via online analysis are still required to complete the IDARS application and RDUa as described above.

For more information about any of NACJD's restricted data access options, please contact [icpsr-help@umich.edu](mailto:icpsr-help@umich.edu).