CC BY NC SA

Restricted Data, Human Subject Review & Secure Data Enclaves

Restricted Data

What Restricted Data Access?

A number of federal agencies and research organizations distribute special files to investigators on which they impose use restrictions. Some individual researchers may collect data containing direct or indirect identifiers which remain in the dataset to enable re-analysis. These files generally contain data fields such as social security numbers, names, geographic detail, or extensive life history markers that might enable an unauthorized user to identify a participant. The use restrictions vary, but they typically involve secure (locked) data storage and password protected computers, and forbid the storage of data on computer hard drives that may be accessed through a computer network connection. These agreements may also limit the types of analyses that are done by the investigator.

How do you obtain restricted data?

To enable use of restricted data most distributors require both an application and a data protection plan. There may also be fees for processing your application; sometimes these are waived for students.

Although there are alternative ways to assure security for the data and applicants should prepare their plans in a manner that best meets their needs, some or all of the following features are typically found in successful data protection plans:

1. Password protection for all files containing data (note that password protection is not regarded as sufficient protection by itself)
2. Removable storage media holding the data (e.g., CDs, diskettes, zip disks, etc.) kept in a locked compartment/room when not in use
3. Printouts derived from data analysis stored in a locked compartment/room when not in use
4. No storage of the data any network, including LANs, Internet enabled, etc.
5. No transmittal of data or analysis output derived from the data via e-mail, e-mail attachments, or FTP (either over the Internet, an Intranet system, or within a local area network)
6. Use of the data on a dedicated computer kept in a secure room and not connected to a network
7. No backup copies of the data to be made
8. Data stored in strongly encrypted form

Here are links to further information on applying to access and some frequently used studies:

AddHealth
Community Tracking Study and other health data
General Social Survey
ICPSR Restricted Access data
National Center for Education Statistics
National Center for Health Statistics

Human Subjects Review

Whenever research involves the collection of data from individuals, a review of how the respondents' privacy and confidentiality will be protected is carried out.

At UCLA, grant applications involving the original collection of data from human subjects must undergo a Human Subjects Review prior to submitting proposals to funding agencies. The Office of the Human Research Protection Program (OHRPP) has a number of policies and guidelines to assist you in ensuring you have addressed data protection issues. The North General Institutional Review Board reviews social-behavioral research, including anthropology, psychology, sociology, education, linguistics, Anderson Graduate School of Management and economics. The South General Institutional Review Board reviews social-behavioral research related to public health, including health services, quality of care, quality of life, health prevention and education, psycho-social drug abuse research, etc. An Investigators Manual is available from OHRPP detailing all aspects of human subjects review.

Human Subjects and Re-use of Public Data

When a researcher uses a publicly available file for study or teaching, this work is referred to as secondary analysis. Publicly available final research data files do not require full human subjects review for their use in secondary analysis. Such files contain no identifiers and are publicly available.

Human Subjects and ICPSR Resources

ICPSR provides resources to assist researchers in satisfying requirements of human subjects review. You can review ICPSR's materials on data protection and human subjects review resources to assist you in working with the UCLA OHRPP. ICPSR also provides guidance on how their data processing methods addresses protection of human subjects.

Additional Resources

American Statistical Association: Privacy and Confidentiality Site
National Science Foundation FAQ

Secure Data Enclave

Some organizations create secure environments where investigators can conduct analyses on restricted data; these environments are referred to as "data enclaves" or "secure data enclaves". Investigators have to file a research plan and meet strict criteria for what can be printed, saved, and removed from the site.

Example of a data enclave at UCLA:

The California Center for Population Research Secure Data Enclave

UCLA investigators planning to access data in a “secure data enclave” must submit an application to the UCLA

IRB/OPRS to obtain UCLA IRB review and approval or Certification of Exemption from UCLA IRB review prior to access to the data.

Enclaves that do not collect, bank, and share data, but rather provide a venue where investigators "work" with specific data sets obtained for their protocol do not require UCLA IRB review and approval or Certification of Exemption from UCLA IRB review.

Secure Data Enclave at ICPSR

ICPSR maintains a Secure Data Enclave and researchers may deposit data into this enclave as well as apply for access to materials residing there. For more details, check the information and guidelines at ICPSR.