

DSDR Preservation Only Policy

DSDR believes that demographic and population study data that have historical or scientific significance should be preserved even if immediate dissemination is not feasible. In such cases, DSDR will accept data for archiving as preservation-only without public and/or restricted use data dissemination. It is important to the scientific and research community that data are archived for long-term preservation to insure that such data are maintained for future research as well as a historical record of current and past research.

The DSDR Preservation Only Policy is designed to address the issues surrounding data preservation without dissemination, to identify the benefits of a Preservation Only Policy and to highlight the situations in which a Preservation Only Policy could be implemented.

Criteria for Applying the Preservation-Only Policy

This list is intended to address and anticipate most circumstances. The criteria will be modified as needed to suit specific situations. DSDR has identified key reasons that provide criteria for data to be archived as preservation only without dissemination. The central focus of preservation only is to provide data producers an independent location to save a back-up copy of their research data.

Content and Scope

DSDR seeks to acquire and preserve data at-risk of being lost to the research community and that are considered key, focal, or important research in the demographic or population studies arena. DSDR selection criteria may include, among others, subject matter, information source, degree of uniqueness or originality, and the nature of the techniques used to represent the information (e.g., physical media, digital media, format) to determine if data can be deposited under Preservation Only Policy.

Complex data or data in uncommon formats that are not amenable to dissemination under current practice are candidates for preservation only.

It is ideal for data to be preserved in communication or cooperation with the original data collectors. The preservation only policy may address this preservation objective by allowing current data (e.g. wave 3 of 10 waves) of ongoing studies (e.g., long-term, longitudinal, or multi-phase studies that cross several years of data collection) to be preserved with the assistance of the current data collection staff.

Disclosure and Confidentiality Issues

Demographic and population data often include sensitive information on respondents, households, neighborhoods, schools, and other institutions. With such data, there is a potential for a disclosure of human subjects and a breach of confidentiality. The data producer and DSDR may wish to preserve the data without dissemination because the data are simply too sensitive for either public or restricted dissemination.

Under preservation only, data producers can deposit restricted data that contain select variables that are simply too sensitive to ever be available for public dissemination. For example, a data producer may disseminate a less sensitive version of the data that has been altered or modified for restricted release, but the original and highly sensitive version is archived with DSDR under preservation only with no dissemination.

Data Producer Dissemination

Data producers may want to maintain control over dissemination and deposit a copy of the data for preservation with DSDR. In such cases, the data producer will need to provide documentation of the dissemination procedures under which the data are provided to the public.

Data Documentation Requirements

Data deposited as preservation only must be fully documented. The DSDR “Guide to Social Science Data Preparation and Archiving” provides a detailed description of data documentation (<http://www.icpsr.umich.edu/DSDR/dsdr-guide.pdf>). Documentation should include funding sources; project description; sample and sampling procedures; weighting; date, geographic location of data collection and time period covered; data sources; and unit(s) of analysis/observation. Other information that should be supplied includes technical information on files; data collection instruments; flowchart of data collection instrument; index or table of contents; list of abbreviations and other conventions; interviewer guide; recode logic; and coding instrument.

Variable documentation should include wording of the question or meaning of the data; question text; missing data codes; imputation or editing information; details on constructed variables; exact meaning of codes; location in data file(s); and variable groupings.

In accordance with the Open Archival, Information System Reference Model (ISO 14721:2003), the community should be able to understand the content information and the preservation description information. In other words, the community should be able to understand the information without needing the assistance of the experts who produced the information.

Data Processing Levels

Preservation only data will undergo minimal data processing. Under minimal processing, the data will be checked for the appropriate frequencies per variable and accurate data locations as specified by the investigator’s codebook. DSDR will perform frequency checks on the data. Identified data discrepancies and other problems, if any, will be documented and archived with the data. Data will be stored in the format received (e.g., SPSS, SAS, STATA, etc.) as well as ASCII.

Given ICPSR’s ongoing commitment to preservation and migrating data formats, DSDR will reserve the choice to update preservation only data to current and evolving formats.

Preservation Only Contract Review

DSDR will not distribute data deposited as preservation only. DSDR will conduct a review with the depositor every 5 years to ascertain the status of the data dissemination. If the existing data dissemination and access channels cease, DSDR will automatically assume dissemination responsibilities as negotiated and outlined in the Preservation Only Contract agreed upon between DSDR and the Data Producer.

Preservation only data will be searchable through ICPSR and users will be directed to the data producer dissemination site to obtain the Preservation Only Policy data.

Enclave Data Preservation

DSDR can archive preservation only copies of highly restricted data that cannot be made available even under restricted data use agreements.