

DSDR

Background

DSDR believes that demographic and population study data that have historical or scientific significance should be preserved even when immediate dissemination is not possible.

Data whose preservation, cataloguing, or storage has been supported by funding agencies may be required to be available at an agreed upon date in the future. However, the data producer may have concerns about having the data immediately released and a delayed dissemination process could provide an approach to address concerns and issues and prepare the data for public release at a future date agreed to by the data producer.

The DSDR delayed dissemination policy (DDP) is designed to address the issues surrounding delayed dissemination, to identify the benefits to a DDP, and to highlight the situations in which a DDP could be implemented.

Data Criteria for Delayed Dissemination

There may be a myriad of reasons as to why data should be archived and preserved without immediate dissemination. DSDR has identified key reasons for data to be archived and preserved for delayed dissemination

Disclosure Risk 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 immediate dissemination until disclosure risk is minimized. For example, smaller longitudinal studies or specific geographic studies (e.g., specific cities identified) may pose disclosure risk even with the removal of direct and indirect identifiers. In this instance, DSDR will work with the data collector to determine when disclosure risk will be minimized and plan to disseminate data at that agreed upon time.

In addition, data collectors can deposit data under the DSDR DDP with an agreement that DSDR conduct a disclosure and confidentiality review to address disclosure risk. The DDP is useful for data with disclosure risk issues when a viable subset of the data can be released through one of DSDR dissemination streams – restricted release and/or secure data enclave. In this situation, the data are archived and preserved in their original form. Working with the data producer, DSDR can provide options under which the data can be disseminated at a later date. For example, as a result of the disclosure review, DSDR can disseminate the data under restricted contract or provide a version of the data that has been altered to minimize disclosure risk (e.g., category collapsing).

Self-Dissemination

Data producers may choose to disseminate their own data, especially while funding is still available to support this activity. DSDR may make a preservation copy during the period of self-dissemination for a number of reasons: 1) to develop expertise with the data, 2) to process the data while there are knowledgeable staff with whom to work, 3) to produce DSDR standardized data, metadata, and study description, and 4) for general safe-keeping. DSDR will enter into an agreement to make an archival copy for later dissemination only when there is an agreed upon date for eventual dissemination. Under the DDP, DSDR can work with the data producer to clarify data issues and develop quality documentation for both DSDR and the data producer. For example, DSDR may request to begin dissemination when grant funding ceases to support dissemination.

Complex, Hierarchical, Longitudinal Data

Data producers may want to archive data that is complex in nature, but wait for public dissemination until all components of the study are complete. For example, producers may want to archive data from a first wave of a longitudinal study with five waves of data collection and want to wait until all five waves of data are collected so that the study data can be disseminated in its entirety.

Under delayed dissemination, DSDR can work with the data producer to structure the data for public dissemination and provide additional documentation, such as variable continuity tables, to make the data user-friendly once it is available for public dissemination.

Other Issues

Occasionally, depositors will request an embargo be placed on their data until a book, article, or a report is written or published. Embargoes will typically last no longer than one year from the time DSDR receives the study. During the embargo, DSDR will process the data in the usual manner, but not release the data to the public. When the embargo is lifted, data would be ready for immediate dissemination.

Delayed Dissemination Release Dates and Timelines

Data will be disseminated at the predetermined end of the preservation-only period. Dissemination dates may vary according to the original reason for non-dissemination. In all cases, the agreed upon date will be negotiated by the data collector and DSDR prior to submission.

Preservation-Only Data under Delayed Dissemination

Issues regarding embargo periods and human subject protections specific to preservation-only studies will be settled as part of the original negotiations with the data collector prior to deposit. In addition, the negotiation process will include ground rules on the extent of processing which will take place by DSDR staff while the study remains in the "preservation-only category".

Possible items to address may include discussion about whether data will be subject to "routine" or "intensive" processing. Highest priority will be given to activities relating to the original reasons for inclusion of the study for "preservation-only". These original reasons include the need to develop expertise with the data and to process the data while there is knowledgeable staff to work with.

Like all data in DSDR, preservation-only studies will be preserved, cataloged (by PI, title, summary, and DSDR-assigned collection/study number) and will be searchable by the tools residing on the DSDR homepage. Unlike other studies, these preservation-only data will not be available for ordering in the normal manner. A new field will be added to the study record - Anticipated Release Date. For data that are being self-disseminated - a link to the website where data can be downloaded/order will appear (as is already customary for self-dissemination datasets).

Summary

The DSDR DDP is intended to provide data producers with another option when considering archiving their demographic and population data. Delayed dissemination is an avenue to encourage data producers to preserve their data with DSDR. Early deposits of data at DSDR allow staff to fully document data files with the help of data program staff. The DSDR DDP program recognizes that many data producers would like to retain control of data dissemination and would like the services of a fully functioning archive.