

CORNELL UNIVERSITY
Cornell Institute for Social and Economic
Research Policy

POLICY
Volume: AC

Responsible Executive:
CISER Data Librarian

Responsible Office:
Cornell Institute for
Social and Economic
Research

Originally Issued:
Revised: 4/3/14

CISER Data Archive Versioning Policy

REASON FOR POLICY

To ensure that data versioning criteria are consistently applied to changes in data files and data documentation in the CISER Data Archive.

POLICY GUIDELINES

For reasons detailed below a significant percentage of data studies are altered within the first year of being published. These changes may be initiated by the data producer, the Data Archive, or data user. CISER has adopted a set of criteria which distinguish between significant and minor changes to a data study to determine whether a new dataset version is merited.

These guidelines ensure that data versioning criteria are consistently applied to changes in data files and data documentation (including correction for error, amendments, additional variables, changes in access conditions, format changes) for inclusion in the CISER Data Archive. This will often involve working closely with the data producer.

Significant changes are those that will have a high impact on the use or interpretation of the data, whereby minor changes are those that will have a low impact in relation to interpretation or use for research purposes. CISER assigns a new version to a data study with significant changes such as: addition of new variables; incorrect data supplied; miscoding of data; formatting changes; substantial documentation changes; changes in access conditions; withdrawal of data elements or documentation files. Minor changes such as small changes in variable labels, spelling corrections in metadata, minor changes in documentation will be made

to relevant content and recorded in ancillary and accompanying documentation. Any and all changes are recorded in a README file.

Where possible CISER will clearly label and make available earlier versions of data and documentation through the data catalog. Version record numbers are captured in metadata held in CISER relational databases. CISER retains the right to withdraw an older version of a data study where significant change may be misrepresentative.

In some cases data studies are issued as completely new study editions (this may be as part of a series) if there are changes to data/variables, major changes to documentation, new waves have been added to a series, or there have been methodological changes.

CISER will cease to hold earlier editions without Cornell provenance or versions of datasets currently held at CISER and also at another archive (such as ICPSR). In the case of ICPSR all earlier editions or versions are available through a back-up system via authorized staff upon request.

ENTITIES AFFECTED BY THIS POLICY

CISER Data Librarian, CISER Users

WHO SHOULD READ THIS POLICY

CISER staff; All users of CISER datasets including students, faculty and researchers at Cornell University.

CONTACTS

If you have questions about specific issues regarding this CISER Data Archive Versioning Policy, call the following offices:

William Block	block@cornell.edu	CRADC Director	607-255-9026
Data Librarian	ciser@cornell.edu	Data Librarian	607-255-4801

RESPONSIBILITIES

The following are major responsibilities each party has in connection with this policy.

CISER Director of Information Technology	Implement operational, physical, and technical equipment and tools to facilitate compliance.
--	--

CISER Data Librarian	
CISER Users	Responsible for understanding obligations and complying with this policy.

Any employee found to have violated this policy may be subject to disciplinary action, up to and including termination of employment.