

DATA MANAGEMENT PLAN

WITH

**THE DATAVERSE
NETWORK**

Mercè Crosas, Ph.D.

Director of Product Development



The Institute
for Quantitative Social Science
at Harvard University

1.

DATA
MANAGEMENT
PLAN (DMP)

2.

DATVERSE
NETWORK
(DVN)

3.

DMP
with
DVN



Engineering

design element

ENG Home

About ENG

Funding Opportunities

Awards

Events

Discoveries

Publications

Advisory Committee

NSF Data Management Plan Requirements

Beginning January 18, 2011, proposals submitted to NSF must include a supplementary document of no more than two pages labeled "Data Management Plan" (DMP). This supplementary document should describe how the proposal will conform to NSF policy on the dissemination and sharing of research results. Proposals that do not include a DMP will not be able to be submitted. For more information about this new requirement, please see the [Grant Proposal Guide, Chapter II.C.2.j](#) and the [Data Management and Sharing Frequently Asked Questions\(FAQs\)](#).

Please note the Engineering Directorate (ENG) has additional guidance for proposals submitted to ENG programs. https://www.nsf.gov/pubs/enggeneral/ENG_DMP_Policy.pdf. Questions and/or suggestions about this new requirement may be addressed to Dr. Mariana Burka at mburka@nsf.gov.

Based on NSF Generic DMP

DATA MANAGEMENT PLAN

CHECKLIST

Data Description**What data will be generated? How will you create the data?
(simulated, observed, experimental, software, physical collections)**

Existing Data

Will you be using existing data? Relationship between the data you are collecting and existing data.

Audience

Who will potentially use the data?

Access and Sharing

How will data files be shared? How will others access them?

Formats

What data formats will you be creating?

Metadata and Documentation

What documentation will you provide to describe the data? Metadata formats and standards?

Storage, backup, replication,
versioning

Are the data files backed up regularly? Are there replicas in different locations? Are older versions of the data kept?

Security

Are the system and storage that will be used secure?

Budget

Any costs for preparing the data? Costs for storage and long-term access?

Privacy, Intellectual Property

Does the data contain private or confidential information? Any copyrights?

Archiving, Preservation, Long-
term Access

What plans do you have to archive the data and other research products? Will it have long-term accessibility?

Adherence

How will you check for adherence of this plan?

Data Description	What data will be generated? How will you create the data? (simulated, observed, experimental, software, physical collections)
Existing Data	Will you be using existing data? Relationship between the data you are collecting and existing data.
Audience	Who will potentially use the data?
Access and Sharing	How will data files be shared? How will others access them?
Formats	What data formats will you be creating?
Metadata and Documentation	What documentation will you provide to describe the data? Metadata formats and standards?
Storage, backup, replication, versioning	Are the data files backed up regularly? Are there replicas in different locations? Are older versions of the data kept?
Security	Are the system and storage that will be used secure?
Budget	Any costs for preparing the data? Costs for storage and long-term access?
Privacy, Intellectual Property	Does the data contain private or confidential information? Any copyrights?
Archiving, Preservation, Long-term Access	What plans do you have to archive the data and other research products? Will it have long-term accessibility?
Adherence	How will you check for adherence of this plan?

Data Description	What data will be generated? How will you create the data? (simulated, observed, experimental, software, physical collections)
Existing Data	Will you be using existing data? Relationship between the data you are collecting and existing data.
Audience	Who will potentially use the data?
Access and Sharing	How will data files be shared? How will others access them?
Formats	What data formats will you be creating?
Metadata and Documentation	What documentation will you provide to describe the data? Metadata formats and standards?
Storage, backup, replication, versioning	Are the data files backed up regularly? Are there replicas in different locations? Are older versions of the data kept?
Security	Are the system and storage that will be used secure?
Budget	Any costs for preparing the data? Costs for storage and long-term access?
Privacy, Intellectual Property	Does the data contain private or confidential information? Any copyrights?
Archiving, Preservation, Long-term Access	What plans do you have to archive the data and other research products? Will it have long-term accessibility?
Adherence	How will you check for adherence of this plan?

Data Description	What data will be generated? How will you create the data? (simulated, observed, experimental, software, physical collections)
Existing Data	Will you be using existing data? Relationship between the data you are collecting and existing data.
Audience	Who will potentially use the data?
Access and Sharing	How will data files be shared? How will others access them?
Formats	What data formats will you be creating?
Metadata and Documentation	What documentation will you provide to describe the data? Metadata formats and standards?
Storage, backup, replication, versioning	Are the data files backed up regularly? Are there replicas in different locations? Are older versions of the data kept?
Security	Are the system and storage that will be used secure?
Budget	Any costs for preparing the data? Costs for storage and long-term access?
Privacy, Intellectual Property	Does the data contain private or confidential information? Any copyrights?
Archiving, Preservation, Long-term Access	What plans do you have to archive the data and other research products? Will it have long-term accessibility?
Adherence	How will you check for adherence of this plan?

Data Description	What data will be generated? How will you create the data? (simulated, observed, experimental, software, physical collections)
Existing Data	Will you be using existing data? Relationship between the data you are collecting and existing data.
Audience	Who will potentially use the data?
Access and Sharing	How will data files be shared? How will others access them?
Formats	What data formats will you be creating?
Metadata and Documentation	What documentation will you provide to describe the data? Metadata formats and standards?
Storage, backup, replication, versioning	Are the data files backed up regularly? Are there replicas in different locations? Are older versions of the data kept?
Security	Are the system and storage that will be used secure?
Budget	Any costs for preparing the data? Costs for storage and long-term access?
Privacy, Intellectual Property	Does the data contain private or confidential information? Any copyrights?
Archiving, Preservation, Long-term Access	What plans do you have to archive the data and other research products? Will it have long-term accessibility?
Adherence	How will you check for adherence of this plan?

Data Description	What data will be generated? How will you create the data? (simulated, observed, experimental, software, physical collections)
Existing Data	Will you be using existing data? Relationship between the data you are collecting and existing data.
Audience	Who will potentially use the data?
Access and Sharing	How will data files be shared? How will others access them?
Formats	What data formats will you be creating?
Metadata and Documentation	What documentation will you provide to describe the data? Metadata formats and standards?
Storage, backup, replication, versioning	Are the data files backed up regularly? Are there replicas in different locations? Are older versions of the data kept?
Security	Are the system and storage that will be used secure?
Budget	Any costs for preparing the data? Costs for storage and long-term access?
Privacy, Intellectual Property	Does the data contain private or confidential information? Any copyrights?
Archiving, Preservation, Long-term Access	What plans do you have to archive the data and other research products? Will it have long-term accessibility?
Adherence	How will you check for adherence of this plan?

Data Description	What data will be generated? How will you create the data? (simulated, observed, experimental, software, physical collections)
Existing Data	Will you be using existing data? Relationship between the data you are collecting and existing data.
Audience	Who will potentially use the data?
Access and Sharing	How will data files be shared? How will others access them?
Formats	What data formats will you be creating?
Metadata and Documentation	What documentation will you provide to describe the data? Metadata formats and standards?
Storage, backup, replication, versioning	Are the data files backed up regularly? Are there replicas in different locations? Are older versions of the data kept?
Security	Are the system and storage that will be used secure?
Budget	Any costs for preparing the data? Costs for storage and long-term access?
Privacy, Intellectual Property	Does the data contain private or confidential information? Any copyrights?
Archiving, Preservation, Long-term Access	What plans do you have to archive the data and other research products? Will it have long-term accessibility?
Adherence	How will you check for adherence of this plan?

Data Description	What data will be generated? How will you create the data? (simulated, observed, experimental, software, physical collections)
Existing Data	Will you be using existing data? Relationship between the data you are collecting and existing data.
Audience	Who will potentially use the data?
Access and Sharing	How will data files be shared? How will others access them?
Formats	What data formats will you be creating?
Metadata and Documentation	What documentation will you provide to describe the data? Metadata formats and standards?
Storage, backup, replication, versioning	Are the data files backed up regularly? Are there replicas in different locations? Are older versions of the data kept?
Security	Are the system and storage that will be used secure?
Budget	Any costs for preparing the data? Costs for storage and long-term access?
Privacy, Intellectual Property	Does the data contain private or confidential information? Any copyrights?
Archiving, Preservation, Long-term Access	What plans do you have to archive the data and other research products? Will it have long-term accessibility?
Adherence	How will you check for adherence of this plan?

Data Description	What data will be generated? How will you create the data? (simulated, observed, experimental, software, physical collections)
Existing Data	Will you be using existing data? Relationship between the data you are collecting and existing data.
Audience	Who will potentially use the data?
Access and Sharing	How will data files be shared? How will others access them?
Formats	What data formats will you be creating?
Metadata and Documentation	What documentation will you provide to describe the data? Metadata formats and standards?
Storage, backup, replication, versioning	Are the data files backed up regularly? Are there replicas in different locations? Are older versions of the data kept?
Security	Are the system and storage that will be used secure?
Budget	Any costs for preparing the data? Costs for storage and long-term access?
Privacy, Intellectual Property	Does the data contain private or confidential information? Any copyrights?
Archiving, Preservation, Long-term Access	What plans do you have to archive the data and other research products? Will it have long-term accessibility?
Adherence	How will you check for adherence of this plan?

Data Description	What data will be generated? How will you create the data? (simulated, observed, experimental, software, physical collections)
Existing Data	Will you be using existing data? Relationship between the data you are collecting and existing data.
Audience	Who will potentially use the data?
Access and Sharing	How will data files be shared? How will others access them?
Formats	What data formats will you be creating?
Metadata and Documentation	What documentation will you provide to describe the data? Metadata formats and standards?
Storage, backup, replication, versioning	Are the data files backed up regularly? Are there replicas in different locations? Are older versions of the data kept?
Security	Are the system and storage that will be used secure?
Budget	Any costs for preparing the data? Costs for storage and long-term access?
Privacy, Intellectual Property	Does the data contain private or confidential information? Any copyrights?
Archiving, Preservation, Long-term Access	What plans do you have to archive the data and other research products? Will it have long-term accessibility?
Adherence	How will you check for adherence of this plan?

Data Description	What data will be generated? How will you create the data? (simulated, observed, experimental, software, physical collections)
Existing Data	Will you be using existing data? Relationship between the data you are collecting and existing data.
Audience	Who will potentially use the data?
Access and Sharing	How will data files be shared? How will others access them?
Formats	What data formats will you be creating?
Metadata and Documentation	What documentation will you provide to describe the data? Metadata formats and standards?
Storage, backup, replication, versioning	Are the data files backed up regularly? Are there replicas in different locations? Are older versions of the data kept?
Security	Are the system and storage that will be used secure?
Budget	Any costs for preparing the data? Costs for storage and long-term access?
Privacy, Intellectual Property	Does the data contain private or confidential information? Any copyrights?
Archiving, Preservation, Long-term Access	What plans do you have to archive the data and other research products? Will it have long-term accessibility?
Adherence	How will you check for adherence of this plan?

Data Description	What data will be generated? How will you create the data? (simulated, observed, experimental, software, physical collections)
Existing Data	Will you be using existing data? Relationship between the data you are collecting and existing data.
Audience	Who will potentially use the data?
Access and Sharing	How will data files be shared? How will others access them?
Formats	What data formats will you be creating?
Metadata and Documentation	What documentation will you provide to describe the data? Metadata formats and standards?
Storage, backup, replication, versioning	Are the data files backed up regularly? Are there replicas in different locations? Are older versions of the data kept?
Security	Are the system and storage that will be used secure?
Budget	Any costs for preparing the data? Costs for storage and long-term access?
Privacy, Intellectual Property	Does the data contain private or confidential information? Any copyrights?
Archiving, Preservation, Long-term Access	What plans do you have to archive the data and other research products? Will it have long-term accessibility?
Adherence	How will you check for adherence of this plan?



Create ready-to-use data management plans for specific funding agencies.

Photo courtesy of Argonne National Laboratory

The DMP Tool allows you to: **1** **2** **3** **4**

**Data Management Plan
Atmospheric CO2 Concentrations,
Mauna Loa Observatory, 2011-2013**

1. Types of data produced

All samples at Mauna Loa Observatory will be collected continuously from air intakes located at five towers – a central tower and four towers located at compass quadrants. Raw data files will contain continuously measured CO2 concentrations, calibration standards, reference standards, daily check standards, and blanks. The sample lines located at compass quadrants were used to examine the influence of source effects associated with wind directions (24). In addition to the CO2 data, we will record weather data (wind speed and direction, temperature, humidity, precipitation, and cloud cover). Site conditions at Mauna Loa Observatory will also be noted and retained.

[See a plan created with the DMP Tool](#)

Recent DMP News

[Video demo now available](#)

[Test Drive the DMP Tool at the ESA](#)

[New requirements from NSF Biological Sciences Directorate](#)

[More news >](#)

DMPTOOL is a service of the [University of California Curation Center](#) of the [California Digital Library](#)
Copyright © 2010-2011 The Regents of the University of California
[Privacy Statement](#) | [Terms of Use](#)

CDL **DMP Tool**
helps you prepare your plan
walking you through similar
QUESTIONS:

1. Types of data produced
2. Data and metadata standards
3. Policies for access and share data
4. Policies for re-use, redistribution
5. Plans for archiving and preservation

OUR RESEARCH PURPOSE

To enable data archiving and preservation through re-formatting, standards and exchange protocols.

To provide control and recognition for data owners through data management and persistent citations.

Dataverse

Create your dedicated web space through a dataverse and promote your brand, publish and share research data and increase scholarly recognition.

Learn more:

- [Publish Data](#)
- [Contributors, Curators and Admins Guide](#)
- [Fact Sheets](#)



Create a Dataverse

Examples of live [dataverses at IQSS](#):

- [Criminall College Libraries](#)
- [Review of Economics and Statistics](#)
- [Murray Research Archive](#)

[Installers Guide](#)

[Data User Guide](#)

[Contributors, Curators, Admin Guide](#)

[Network Administrators Guide](#)

[Developers Guide](#)

[Reference Materials](#)

[Version 2.1.2 Downloads Now Available on Sourceforge](#)

[Yale Law School Roundtable Recommends the Dataverse Network](#)

[Version 2.1.1 Downloads Now Available on Sourceforge](#)

[Version 2.1.1 Rolled Out to IQSS Dataverse Network](#)

DATAVERSE NETWORK

AN OVERVIEW

WHY

the

Dataverse Network?

A PROBLEM

Different needs for Archives, Data Libraries, Researchers, Journals, Funding Agencies...

We should preserve the data!

Who will curate the data?

We need persistent links!

I want recognition for my data!

I want to review submissions and control who has access

A SOLUTION

Centralized Professional Archiving with
Distributed Control and Recognition

- Persistent identifiers
- Fixity
- Backups and recovery
- Metadata standards
- Conversion standards
- Preservation standards

+

- Branding and visibility
- Data discovery
- Ease of use
- Scholarly citation
- Control over updates
- Terms of access and use

HOW

does it work?

DATAVERSE
NETWORK

A **centralized** software
installation and data repository

Dataverse

Dataverse

Dataverse

Dataverse

Individual
distributed
data archive
with its own
branding

Collections

Study

Study

Study

Metadata

Data Files

RICH DATA SUPPORT

Data Files

Tabular Data:

STATA

SPSS

CSV + control card

Tab delimited + DDI

Social Network Data:

GraphML

Other Data or Complementary Files:

ALL formats are accepted
BUT only tabular files have
full data support

- Convert to a preservation format (primary data + variable metadata)
- Calculate Universal Numerical Fingerprint (UNF)
- Download in multiple formats
- Download a subset of the data
- Generate summary statistics
- Apply Zelig (R) statistical methods
- Visualize time series
- **Define Terms of Use and Permissions**

FORMAL, PERSISTENT DATA CITATIONS

Authors

Year

Title

Persistent URL and ID

UNF

Distributor

Version

+ other optional fields (e.g., data subset)

Sidney Verba; Joseph DiNunzio; Christina Spaulding, 1985, "Harvard Sexual Harassment Survey, 1983", <http://hdl.handle.net/1902.1/00686>

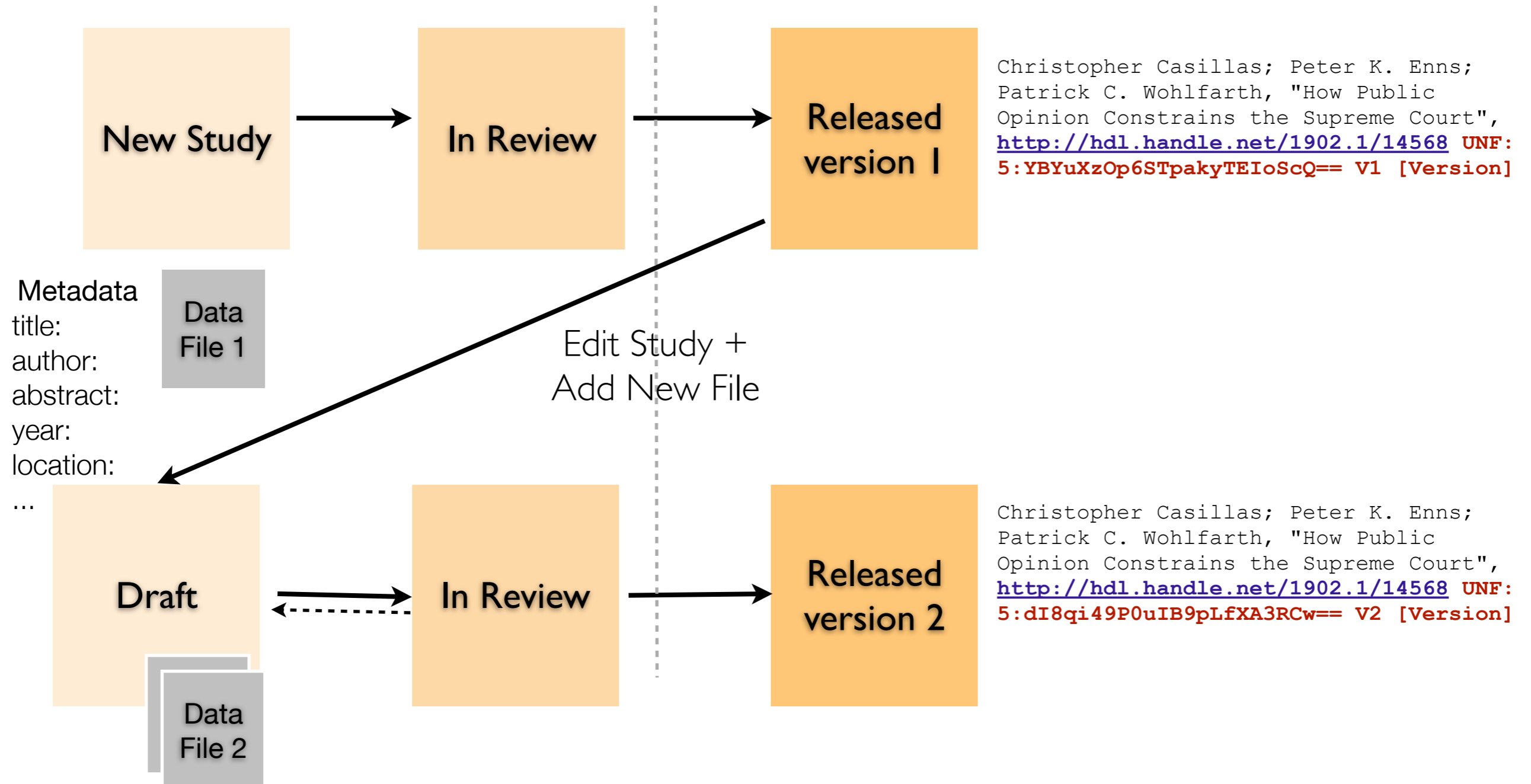
UNF:3:eG6QJsi2VsgIMV49FtGcDA== Murray Research Archive

[Distributor] V1 [Version]

DATA VERSIONING AND MANAGEMENT

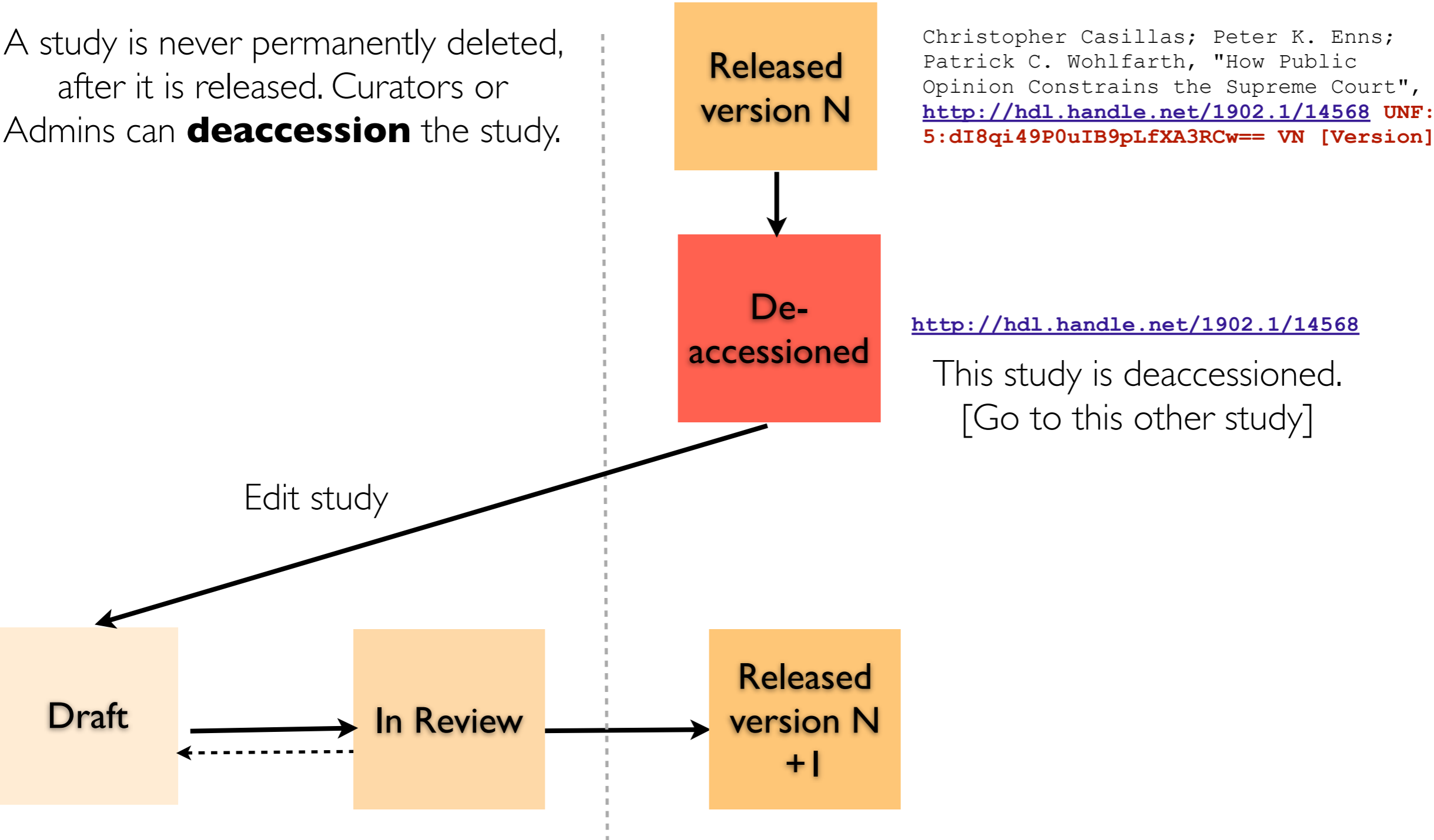
Contributors, curators, admins
view

End user
view



DATA NEVER DELETED

A study is never permanently deleted, after it is released. Curators or Admins can **deaccession** the study.



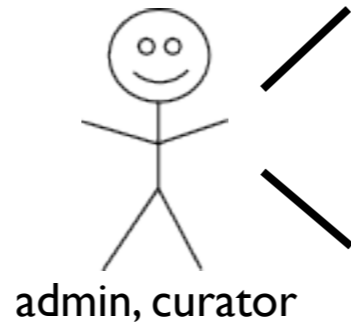
Christopher Casillas; Peter K. Enns; Patrick C. Wohlfarth, "How Public Opinion Constrains the Supreme Court", <http://hdl.handle.net/1902.1/14568> UNF:5:dI8qi49P0uIB9pLfXA3RCw== VN [Version]

<http://hdl.handle.net/1902.1/14568>

This study is deaccessioned. [Go to this other study]

WORKFLOWS FOR ALL

Regular Dataverse

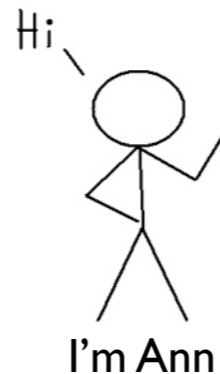


Option A - The **admin or curator** adds Mike as a contributor. Mike can now add studies and edit his **own** studies in this Dataverse.

or

Option B - The **admin or curator** adds Mike as a contributor. Mike can add studies and edit **any** study in this Dataverse.

Open Dataverse

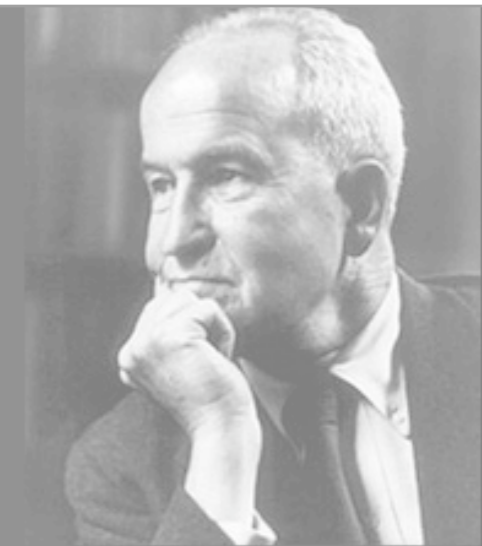


Ann **creates an account** and starts adding and editing her **own** studies in an open Dataverse. The admin or curator reviews the studies before releasing them.

Wiki Dataverse



Chris **creates an account** and starts adding new studies and editing **any** study in a wiki Dataverse. The admin or curator reviews the studies before releasing them.



All IQSS Dataverses >

Murray Research Archive Original Collection Dataverse

Search User Guides Report Issue

POWERED BY THE **Dataverse Network**™ PROJECT

Log In Create Account

The Henry A. Murray Research Archive is the endowed, permanent repository for quantitative and qualitative research data at the Institute for Quantitative Social Science, and provides physical storage for the entire IQSS Dataverse Network. Our collection comprises over 125 terabytes of data, audio, and video.

More information about the Murray Archive can be found at our Website

REGULAR DATAVERSE

Henry A. Murray Research Archive Original Collection

Search Studies [Tips](#)

within this collection

Sort By:

Studies: 501

Attitudes of Participants at the Iowa Women's Political Caucus, 1973 by Barbara Burrell Abstract: This study was designed to obtain a profile of women who are active in a movement generating social change through advocating political equality. The sample consists of participants at the first state ...	hdl:1902.1/00508 112 downloads Last Released: Jun 21, 2011
Survey of Child-Care Needs of Lower-Paid Personnel, 1980 by Harvard-Radcliffe Child Care Council Abstract: The purpose of this study was to identify unmet needs for child care among lower-paid personnel employed at Harvard University. This identification was aimed both at estimating the numbers of lower-paid ...	hdl:1902.1/00566 31 downloads + analyses Last Released: Jun 21, 2011
Early Head Start Research and Evaluation Project, 1996 - 2001 by Administration for Children and Families	hdl:1902.1/00097 2436 downloads + analyses

The Review of Economics and Statistics

All IQSS Dataverses >

Review of Economics and Statistics Dataverse

[Search](#) [User Guides](#) [Report Issue](#)

POWERED BY THE **Dataverse Network**™ PROJECT

[Log In](#) [Create Account](#)

This page contains replication data and code for the articles published in the Review of Economics and Statistics. Please notice that most articles in volumes previous to volume 92 were not subject to the journal's data availability policy.

OPEN DATAVERSE

Create an account to add your own study to this dataverse. Already have an account? [Log In](#)

Review of Economics and Statistics

- Previous Volumes
- Volume 92
 - Volume 92, Issue 1
 - Volume 92, Issue 2
 - Volume 92, Issue 3
 - Volume 92, Issue 4
- Volume 93
 - Volume 93, Issue 1
 - Volume 93, Issue 2
 - Volume 93, Issue 3
 - Volume 93, Issue 4

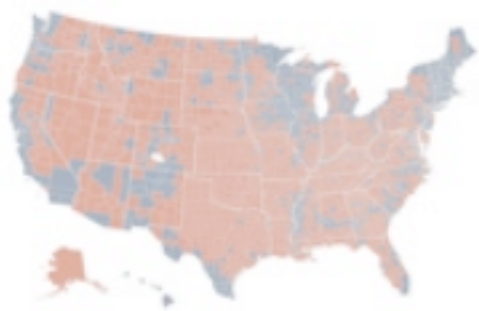
Search Studies [Advanced Search](#)
[Tips](#)

Studies with the collection

1 2 3 4 > >> >>>

Replication data for: New Eurocoin: Tracking Economic Growth in Real Time by Filippo Altissimo; Riccardo Cristadoro; Mario Forni; Marco Lippi; Giovanni Veronese	hdl:1902.1/15447 27 downloads Last Released: Apr 19, 2011
Replication data for: Inequality and Redistribution: Evidence from U.S. Counties and States, 1890-1930 by Rodney Ramcharan	hdl:1902.1/15784 236 downloads Last Released: Apr 19, 2011
Replication data for: Corruption and Bilateral Trade Flows: Extortion or Evasion? by Pushan Dutt; Daniel Traca	hdl:1902.1/15394 28 downloads + analyses Last Released: Mar 31, 2011
Replication data for: Universal Vouchers and Racial and Ethnic Segregation by Eric Brunner; Jennifer Imazeki; Stephen Ross	hdl:1902.1/15356 36 downloads + analyses Last Released: Mar 31, 2011

OPEN DATAVERSE



HARVARD ELECTION DATA ARCHIVE

Sharing and Improving Election Data



HOME DATA PEOPLE MAPS PUBLICATIONS

Data

[All IQSS Dataverses >](#)

Election Data Archive Dataverse

POWERED BY THE **Dataverse Network™** PROJECT

[Search](#) [User Guides](#) [Report Issue](#)

[Log In](#) [Harvard Affiliate](#) [Create Account](#)

WIKI DATAVERSE

The Harvard Election Data Archive is a space to share and improve election data. You can create an account and either request any cataloging information for the studies in the dataverse or upload new data files. Thank you to our colleagues Michael McDonald and David Brudler for their help with data collection.

WIKI DATAVERSE
Create an account to add your own study or contribute to existing studies in the dataverse. Already have an account? [Log In](#)

The Harvard Election Data Archive by [Stephen Ansolabehere](#) and [Jonathan Rodden](#) is licensed under a [Creative Commons Attribution 3.0 Unported License](#).

Election Data Archive

[Advanced Search Tips](#)

Sort By:

Studies: **24**

[Pennsylvania Data Files](#)

by Stephen Ansolabehere; Jonathan Rodden

Abstract: Pennsylvania Shapefile and Election Results

hdl:1902.1/16389

15 downloads + analyses

Last Released: Aug 1, 2011

[Arizona Data Files](#)

by Stephen Ansolabehere; Jonathan Rodden

Abstract: Arizona Shapefile and Election Results

hdl:1902.1/16388

5 downloads + analyses

Last Released: Aug 1, 2011



Astronomy Dataverse Network



[Search](#) [User Guides](#) [Report Issue](#)

[Log Out](#) [Network Admin](#) [Options](#)

This is a prototype for an Astronomy data repository. The project is a collaboration of the Harvard-Smithsonian Center for Astrophysics, the ADS, the Wolbach Library, and IAS, with support from the NASA Science Research Center.

DATAVERSE

All Dataverses

[Advanced Search](#)
[Tips](#)

BEYOND

ALL A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

Dataverses: **3** | Studies: **39** | Files: **253**

SOCIAL SCIENCE

Name	View Info [+]	Affiliation	Released	Activity
ChaMPlane	View Info [+]	Harvard-Smithsonian Center for Astrophysics	Jun 23, 2011	■■■■■
COMPLETE	View Info [+]	Harvard-Smithsonian Center for Astrophysics	Jun 23, 2011	■■■■■
1.2 Meter CO Survey	View Info [+]	Smithsonian Astrophysical Observatory	May 23, 2011	■■■■■





How about Biomedicine, Earth,
Sciences, Physics ...?

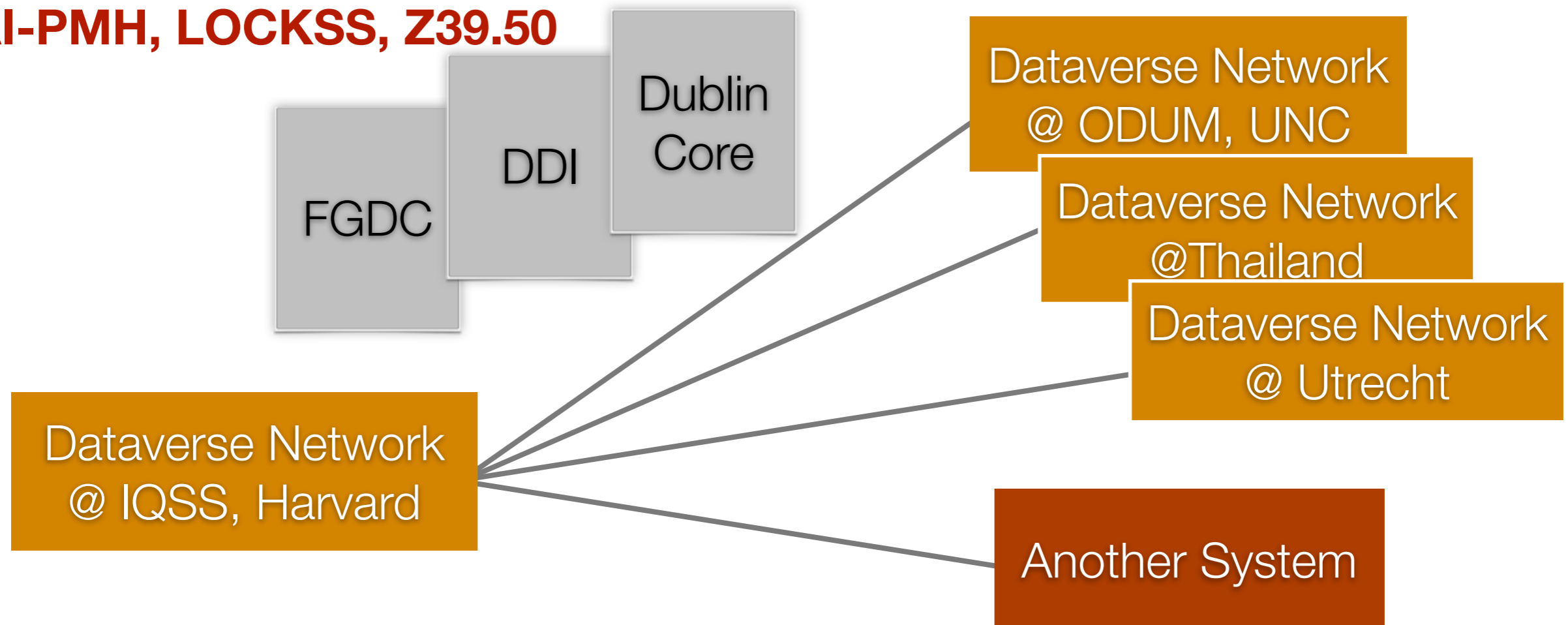
Dataverse is cross-disciplinary.

**We are expanding the study
metadata and building a
community of interested groups.**

Contact us if you are interested.

SUPPORTING STANDARDS

OAI-PMH, LOCKSS, Z39.50



- Study and variable metadata are exported into **XML** (Dublin Core, Data Descriptive Initiative - DDI, ...) and MARC.
- **OAI-PMH** for harvesting metadata.
- **LOCKSS** for data duplication in multiple locations.
- **Z39.50** for distributed search.
- **E-Z Proxy** to authenticate for data access.



NSF Web Site

Engineering

design element

- ENG Home
- About ENG
- Funding Opportunities
- Awards
- News
- Events
- Discoveries
- Publications
- Advisory Committee

NSF Data Management Plan Requirements



An Open-Source Application for Publishing, Citing and Discovering Research Data

- LEARN ABOUT THE PROJECT
- USE AND SHARE DATA
- INSTALL THE SOFTWARE
- CONTRIBUTE TO DEVELOPMENT

OUR RESEARCH PURPOSE

To enable data archiving and preservation through re-formatting, standards and exchange protocols.

To provide control and recognition for data owners through data management and persistent citations.

3. DMP with DVN

Dataverse

Create your dedicated web space through a dataverse and promote your brand, publish and share research data and increase scholarly recognition.

Learn more:

- Publish Data
- Contributors, Curators and Admins Guide
- Fact Sheets

Create a Dataverse

Examples of live dataverses at IQSS:

- Grinnell College Libraries
- Review of Economics and Statistics
- Murray Research Archive

- Installers Guide
- Data User Guide
- Contributors, Curators, Admin Guide
- Network Administrators Guide
- Developers Guide
- Reference Materials

Version 2.1.2 Downloads Now Available on Sourceforge

Yale Law School Roundtable Recommends the Dataverse Network

Version 2.1.1 Downloads Now Available on Sourceforge

Version 2.1.1 Rolled Out to IQSS Dataverse Network

Data Description	What data will be generated? How will you create the data? (simulated, observed, experimental, software, physical collections)
Existing Data	Will you be using existing data? Relationship between the data you are collecting and existing data.
Audience	Who will potentially use the data?
Access and Sharing	How will data files be shared? How will others access them?
Formats	What data formats will you be creating?
Metadata and Documentation	What documentation will you provide to describe the data? Metadata formats and standards?
Storage, backup, replication, versioning	Are the data files backed up regularly? Are there replicas in different locations? Are older versions of the data kept?
Security	Are the system and storage that will be used secure?
Budget	Any costs for preparing the data? Costs for storage and long-term access?
Privacy, Intellectual Property	Does the data contain private or confidential information? Any copyrights?
Archiving, Preservation, Long-term Access	What plans do you have to archive the data and other research products? Will it have long-term accessibility?
Adherence	How will you check for adherence of this plan?

Data Description

**What data will be generated? How will you create the data?
(simulated, observed, experimental, software, physical collections)**

Existing Data

**Will you be using existing data? Relationship between the data you
are collecting and existing data.**

Audience

Who will potentially use the data?

My Data Management Plan

Data collected by this project will include ... [I can answer the first three items on my own] ...

Privacy, Intellectual Property

Does the data contain private or confidential information? Any copyrights?

Archiving, Preservation, Long-term Access

What plans do you have to archive the data and other research products? Will it have long-term accessibility?

Adherence

How will you check for adherence of this plan?

Data Description	What data will be generated? How will you create the data? (simulated, observed, experimental, software, physical collections)
Existing Data	Will you be using existing data? Relationship between the data you are collecting and existing data.
Audience	Who will potentially use the data?
Access and Sharing	How will data files be shared? How will others access them?
Formats	What data formats will you be creating?
Metadata and Documentation	What documentation will you provide to describe the data? Metadata formats and standards?
Storage, backup, replication, versioning	Are the data files backed up regularly? Are there replicas in different locations? Are older versions of the data kept?
Security	Are the system and storage that will be used secure?
Budget	Any costs for preparing the data? Costs for storage and long-term access?
Privacy, Intellectual Property	Does the data contain private or confidential information? Any copyrights?
Archiving, Preservation, Long-term Access	What plans do you have to archive the data and other research products? Will it have long-term accessibility?
Adherence	How will you check for adherence of this plan?



James M. Snyder, Jr.

Professor of Government

1737 Cambridge St. Cambridge, MA 02138, 617-496-1089, jsnyder@gov.harvard.edu (email)



-
- [Bio](#)
- [Class](#)
- [Publications](#)
- [Working Papers](#)
- [Book Chapters](#)

Dataverse

[All IQSS Dataverses >](#)

James Snyder Dataverse

POWERED BY THE **Dataverse Network™** PROJECT

[Search](#) [User Guides](#) [Report Issue](#)

[Log Out](#) [IQSS Admin](#) [Options](#)

James Snyder [Advanced Search Tips](#)

Sort By: Studies: **3**

<p>Replication data for: Electoral Selection with Parties and Primaries</p> <p>by James M. Snyder, Jr.; Michael M. Ting</p>	<p>hdl:1902.1/15624</p> <p> 35 downloads + analyses</p> <p>Last Released: Apr 11, 2011</p>
---	--

- All data collected or generated will be deposited in my Dataverse.
- The Dataverse facilitates data access through search, browse, archival formats, cataloging, permissions, ...

Data Description	What data will be generated? How will you create the data? (simulated, observed, experimental, software, physical collections)
Existing Data	Will you be using existing data? Relationship between the data you are collecting and existing data.
Audience	Who will potentially use the data?
Access and Sharing	How will data files be shared? How will others access them?
Formats	What data formats will you be creating?
Metadata and Documentation	What documentation will you provide to describe the data? Metadata formats and standards?
Storage, backup, replication, versioning	Are the data files backed up regularly? Are there replicas in different locations? Are older versions of the data kept?
Security	Are the system and storage that will be used secure?
Budget	Any costs for preparing the data? Costs for storage and long-term access?
Privacy, Intellectual Property	Does the data contain private or confidential information? Any copyrights?
Archiving, Preservation, Long-term Access	What plans do you have to archive the data and other research products? Will it have long-term accessibility?
Adherence	How will you check for adherence of this plan?



James M. Snyder, Jr.

Professor of Government

1737 Cambridge St. Cambridge, MA 02138, 617-496-1089, jsnyder@gov.harvard.edu (email)



Dataverse

[All IQSS Dataverses >](#)

James Snyder Dataverse

POWERED BY THE **Dataverse Network** PROJECT

[Search](#) [User Guides](#) [Report Issue](#)

[Log Out](#) [IQSS Admin](#) [Options](#)

Bio

- Quantitative data will be deposited either in SPSS, Stata, CSV, Tab delimited, or GraphML.
- Images in JPEG-2000 or TIFF. [good practice]
- Audio in MP3 or WAVE. [good practice]

UPLOAD FILES

✓ Choose a Data Type

Tabular Data

SPSS/POR

SPSS/SAV

Stata

CSV (w/SPSS card)

TAB (w/DDI)

Network Data

GraphML

Other

Tabular and Network Data files can be subset and analyzed using the Dataverse Network analysis tools. These files will take longer to upload and you'll get a notification once the upload is completed. Tabular files will also get Universal Numerical Fingerprint (UNF). All other files types will be available for download only.

Category	File Name	Description	Size (bytes)	Remove
----------	-----------	-------------	--------------	--------

Save

Cancel

Data Description	What data will be generated? How will you create the data? (simulated, observed, experimental, software, physical collections)
Existing Data	Will you be using existing data? Relationship between the data you are collecting and existing data.
Audience	Who will potentially use the data?
Access and Sharing	How will data files be shared? How will others access them?
Formats	What data formats will you be creating?
Metadata and Documentation	What documentation will you provide to describe the data? Metadata formats and standards?
Storage, backup, replication, versioning	Are the data files backed up regularly? Are there replicas in different locations? Are older versions of the data kept?
Security	Are the system and storage that will be used secure?
Budget	Any costs for preparing the data? Costs for storage and long-term access?
Privacy, Intellectual Property	Does the data contain private or confidential information? Any copyrights?
Archiving, Preservation, Long-term Access	What plans do you have to archive the data and other research products? Will it have long-term accessibility?
Adherence	How will you check for adherence of this plan?



- Study metadata will be entered in the Dataverse Cataloging Information page with about 100 fields to choose from.
- Additional documentation will be uploaded in PDF or plain text formats.
- A formal Data Citation will be generated automatically.
- Metadata will be exported in XML (DDI, DC) .

Snyder, JM, Ting M. 2011. "Electoral Selection with Parties and Primaries". American Journal of Political Science. Forthcoming.

Citation Information

Study ID hdl:1902.1/15624

Author *	Name	Affiliation	
	<input type="text" value="James M. Snyder, Jr."/>	<input type="text" value="Harvard University, Dept. of Government"/>	+ -
	(FirstName LastName)		

Author *	Name	Affiliation	
	<input type="text" value="Michael M. Ting"/>	<input type="text" value="Department of Political Science and SIPA Columbi."/>	+ -
	(FirstName LastName)		

Producer *	Producer	
	<input type="text"/>	+

Data Description	What data will be generated? How will you create the data? (simulated, observed, experimental, software, physical collections)
Existing Data	Will you be using existing data? Relationship between the data you are collecting and existing data.
Audience	Who will potentially use the data?
Access and Sharing	How will data files be shared? How will others access them?
Formats	What data formats will you be creating?
Metadata and Documentation	What documentation will you provide to describe the data? Metadata formats and standards?
Storage, backup, replication, versioning	Are the data files backed up regularly? Are there replicas in different locations? Are older versions of the data kept?
Security	Are the system and storage that will be used secure?
Budget	Any costs for preparing the data? Costs for storage and long-term access?
Privacy, Intellectual Property	Does the data contain private or confidential information? Any copyrights?
Archiving, Preservation, Long-term Access	What plans do you have to archive the data and other research products? Will it have long-term accessibility?
Adherence	How will you check for adherence of this plan?



James M. Snyder, Jr.
Professor of Government

1727 Cambridge St. Cambridge, MA 02138 617-495-1080 jsnyder@gov.harvard.edu (email)



- The Dataverse will keep multiple versions of my data.
- Deposited data will never be destroyed (unless legally required).
- In addition, the Dataverse Network at IQSS provides system backups in a daily basis.

Book Chapters

Dataverse

+ Cataloging Information	+ Documentation, Data and Analysis	+ User Comments	- Versions		
Version History					
Difference	Version	Status	Comments	Released	Contributors
<input type="checkbox"/>	3	Released		Mon Apr 11 15:12:13 EDT 2011	IQSS Admin
<input type="checkbox"/>	2	Archived		Tue Apr 05 18:29:06 EDT 2011	IQSS Admin
<input type="checkbox"/>	1	Archived		Tue Mar 08 09:11:00 EST 2011	Helen Lewis

Data Description	What data will be generated? How will you create the data? (simulated, observed, experimental, software, physical collections)
Existing Data	Will you be using existing data? Relationship between the data you are collecting and existing data.
Audience	Who will potentially use the data?
Access and Sharing	How will data files be shared? How will others access them?
Formats	What data formats will you be creating?
Metadata and Documentation	What documentation will you provide to describe the data? Metadata formats and standards?
Storage, backup, replication, versioning	Are the data files backed up regularly? Are there replicas in different locations? Are older versions of the data kept?
Security	Are the system and storage that will be used secure?
Budget	Any costs for preparing the data? Costs for storage and long-term access?
Privacy, Intellectual Property	Does the data contain private or confidential information? Any copyrights?
Archiving, Preservation, Long-term Access	What plans do you have to archive the data and other research products? Will it have long-term accessibility?
Adherence	How will you check for adherence of this plan?



developing
SECURE
technology
SOLUTIONS



ENTERPRISE SECURITY POLICY

RESEARCH DATA SECURITY POLICY

Welcome

- The Dataverse Network at IQSS follows “good computer use practices” set by the Security & Privacy group at Harvard.

CONTACT US

Jump

Go

requirements and policies that support these commitments and works to communicate the policies to the University community. Each Harvard School is responsible for implementing these University guidelines and for developing local policies, where needed, to facilitate a secure environment that is consistent with University requirements.

Harvard community members-- student, staff, and faculty-- encounter sensitive information every day, information such as student grades and evaluations, staff evaluations, credit card numbers, bank accounts, salaries, and personal information including home addresses for example. This information is considered confidential by the University and by the person the information is about.

Data Description	What data will be generated? How will you create the data? (simulated, observed, experimental, software, physical collections)
Existing Data	Will you be using existing data? Relationship between the data you are collecting and existing data.
Audience	Who will potentially use the data?
Access and Sharing	How will data files be shared? How will others access them?
Formats	What data formats will you be creating?
Metadata and Documentation	What documentation will you provide to describe the data? Metadata formats and standards?
Storage, backup, replication, versioning	Are the data files backed up regularly? Are there replicas in different locations? Are older versions of the data kept?
Security	Are the system and storage that will be used secure?
Budget	Any costs for preparing the data? Costs for storage and long-term access?
Privacy, Intellectual Property	Does the data contain private or confidential information? Any copyrights?
Archiving, Preservation, Long-term Access	What plans do you have to archive the data and other research products? Will it have long-term accessibility?
Adherence	How will you check for adherence of this plan?



IQSS Dataverse Network

[Search](#) [Manifest](#) [User Guides](#) [Report Issue](#)



[Log Out](#) [IQSS Admin](#) [Options](#)

Welcome to the IQSS Dataverse Network!

Access the world's largest collection of social science research data here by searching across or browsing through one of the virtual data archives (called "dataverses") listed below. You may also create a dataverse

CREATE A DATAVERSE

Create a Dataverse to upload your own data sets and create collections of data.


- The Dataverse Network at IQSS is free and open to all social science research data.
- Restrictions are 2GB per file, with no limit in the number of files.

(In the future, a fee might apply to archive very large collections - currently under review)

Research Projects
Scholars

panu	View Info [+]	utcc	Aug 10, 2011	■■■■■
Recession Trends	View Info [+]	Stanford University's Center for the Study of Poverty and Inequality	Aug 9, 2011	■ ■■■■
Iyer, Sundeep		Harvard University	Aug 5, 2011	■ ■■■■
Politics of Genomics	View Info [+]	Harvard University	Aug 3, 2011	■■■■■

Data Description	What data will be generated? How will you create the data? (simulated, observed, experimental, software, physical collections)
Existing Data	Will you be using existing data? Relationship between the data you are collecting and existing data.
Audience	Who will potentially use the data?
Access and Sharing	How will data files be shared? How will others access them?
Formats	What data formats will you be creating?
Metadata and Documentation	What documentation will you provide to describe the data? Metadata formats and standards?
Storage, backup, replication, versioning	Are the data files backed up regularly? Are there replicas in different locations? Are older versions of the data kept?
Security	Are the system and storage that will be used secure?
Budget	Any costs for preparing the data? Costs for storage and long-term access?
Privacy, Intellectual Property	Does the data contain private or confidential information? Any copyrights?
Archiving, Preservation, Long-term Access	What plans do you have to archive the data and other research products? Will it have long-term accessibility?
Adherence	How will you check for adherence of this plan?



James M. Snyder, Jr.

- Data uploaded to the Dataverse cannot contain confidential information.
- I agree to the terms of use, which state that the data must be anonymized following IRB-approved protocols, and do not infringe upon the copyrights.

Working Papers

Book Chapters

Dataverse

the "I AGREE" button below, you agree that these Terms and Conditions apply to all Materials that you post to the Service:

1. You represent and warrant to IQSS that you are lawfully entitled and have full authority to license IQSS and the Archive to use the Materials in the ways described in these Terms and Conditions and that you are not under any obligation or restriction created by law, contract or otherwise that would prevent you from entering into and fully performing these Terms and Conditions;
2. You give permission and any required licenses to IQSS and the Archive to store and backup the materials;
3. The Materials do not violate any laws, including but not limited to laws related to defamation or obscenity;
4. The Materials do not infringe upon the copyrights or other intellectual property rights, including, but not limited to

I agree and accept these terms of use.

Continue

Data Description	What data will be generated? How will you create the data? (simulated, observed, experimental, software, physical collections)
Existing Data	Will you be using existing data? Relationship between the data you are collecting and existing data.
Audience	Who will potentially use the data?
Access and Sharing	How will data files be shared? How will others access them?
Formats	What data formats will you be creating?
Metadata and Documentation	What documentation will you provide to describe the data? Metadata formats and standards?
Storage, backup, replication, versioning	Are the data files backed up regularly? Are there replicas in different locations? Are older versions of the data kept?
Security	Are the system and storage that will be used secure?
Budget	Any costs for preparing the data? Costs for storage and long-term access?
Privacy, Intellectual Property	Does the data contain private or confidential information? Any copyrights?
Archiving, Preservation, Long-term Access	What plans do you have to archive the data and other research products? Will it have long-term accessibility?
Adherence	How will you check for adherence of this plan?



James M. Snyder, Jr.
Professor of Government

1737 Cambridge St. Cambridge, MA 02138, 617-496-1089, jsnyder@gov.harvard.edu (email)



- The IQSS Dataverse Network commits to good archival practices.
- Data files are re-formatted for long term access, and replicated to multiple locations through LOCKSS.

Book Chapters

Dataverse

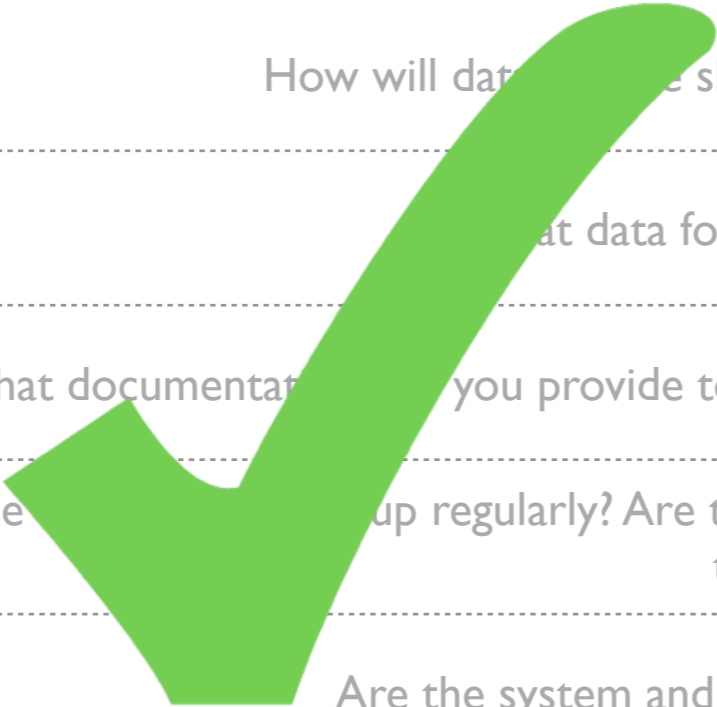
+ Cataloging Information | Documentation, Data and Analysis | + User Comments | + Versions

i Use the check boxes next to the file name to download multiple files. Data files will be downloaded in their default format. You can also download all the files in a category by checking the box next to the category name. You will be prompted to save a single archive file. Study files that have restricted access will not be downloaded.

Select all files Download All Selected Files

<input type="checkbox"/>			
<input type="checkbox"/>	replication.codebook.pdf Adobe PDF - 33 KB - 20 downloads	Download	This codebook describes the variables in the following datasets: tables 1 2 replication.dta table 3 replication.dta
<input type="checkbox"/>	table_3_replication.tab Tab Separated - 601 KB - 9 downloads + analyses	Download as... Tab Delimited Original File Splus Stata R	replication data tables View Data Citation (+)
<input type="checkbox"/>	tables_1_2_replication.tab Tab Separated - 2 KB - 6 downloads + analyses	Download as...	replication data tables View Data Citation (+)
	TABULAR DATA 16156 Cases 9 Variables	Access Subset/Analysis	
	TABULAR DATA 48 Cases 6 Variables		

Data Description	What data will be generated? How will you create the data? (simulated, observed, experimental, software, physical collections)
Existing Data	Will you be using existing data? Relationship between the data you are collecting and existing data.
Audience	Who will potentially use the data?
Access and Sharing	How will data be shared? How will others access them?
Formats	What data formats will you be creating?
Metadata and Documentation	What documentation you provide to describe the data? Metadata formats and standards?
Storage, backup, replication, versioning	Are the data backup regularly? Are there replicas in different locations? Are older versions of the data kept?
Security	Are the system and storage that will be used secure?
Budget	Any costs for preparing the data? Costs for storage and long-term access?
Privacy, Intellectual Property	Does the data contain private or confidential information? Any copyrights?
Archiving, Preservation, Long-term Access	What plans do you have to archive the data and other research products? Will it have long-term accessibility?
Adherence	How will you check for adherence of this plan?



THANKS!

<http://thedata.org>

<http://dvn.iq.harvard.edu>

dvn_support@help.hmdc.harvard.edu

mcrosas@hmdc.harvard.edu

Dataverse Network team and IQSS Archival support: Micah Altman, Leonis Andreev, Sonia Barbosa, Kevin Condon, Gustavo Durand, Michael Heppler, Ellen Kraffmiller, Stephen Kraffmiller, Dwayne Liburd, Robert Treacy