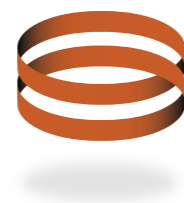


The Dataverse Network

An open-source Application for Sharing,
Discovering and Preserving Data

Mercè Crosas, Ph.D.
Director of Product Development
IQSS, Harvard University

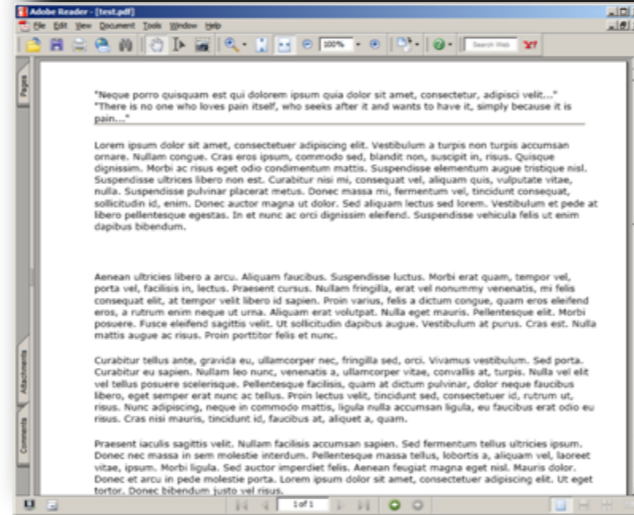


The Institute
for Quantitative Social Science
at Harvard University

Problems with traditional ways to store research data

x0	x1	x2	x3	x4	x5
1	1.06	9.2	151	54.4	1.6
2	0.89	10.3	202	57.9	2.2
3	1.43	15.4	113	53	3.4
4	1.02	11.2	168	56	0.3
5	1.49	8.8	192	51.2	1
6	1.32	13.5	111	60	-2.2
7	1.22	12.2	175	67.6	2.2
8	1.1	9.2	245	57	3.3
9	1.34	13	168	60.4	7.2
10	1.12	12.4	197	53	2.7
11	0.75	7.5	173	51.5	6.5
12	1.13	10.9	178	62	3.7
13	1.15	12.7	199	53.7	6.4
14	1.09	12	96	49.8	1.4
15	0.96	7.6	164	62.2	-0.1
16	1.16	9.9	252	56	9.2
17	0.76	6.4	136	61.9	9
18	1.05	12.6	150	56.7	2.7
19	1.16	11.7	104	54	-2.1
20	1.2	11.8	148	59.9	3.5
21	1.04	8.6	204	61	3.5
22	1.07	9.3	174	54.3	5.9

Data sets

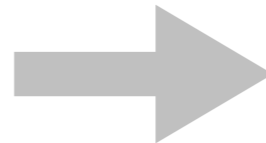


Text



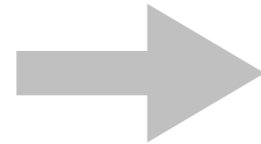
Images

Keep it in my laptop



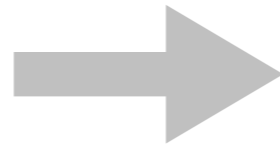
Hard to share, not safe

Send it to an archive



Author doesn't get enough credit

Post it in my web site



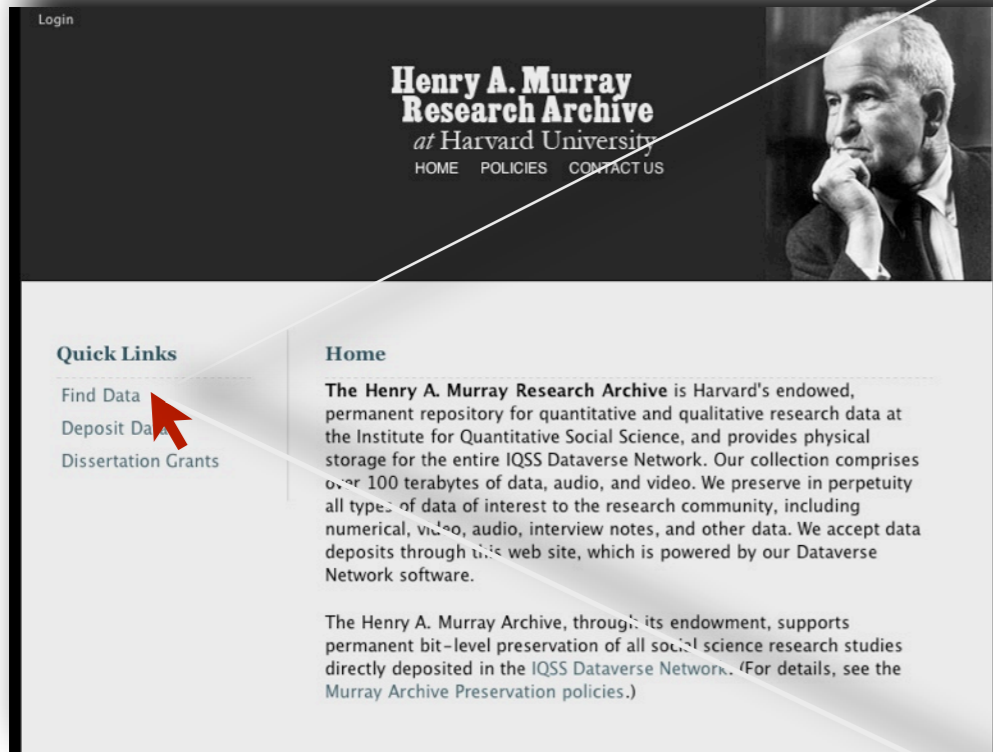
No professional archiving services
(backups, recovering, cataloging,
preservation, etc)

Solving the problems through technology

We've developed a system that:

- ☑ Gives **ownership and recognition** to author, distributor and publisher
- ☑ Generates a **persistent data citation**
- ☑ Converts data sets to a **preservable** and verifiable format
- ☑ Distributes data to the **public**, but also supports **restricted** access
- ☑ Indexes all metadata for quick data **discovery**
- ☑ Supports **subsetting and analysis** for (some) data files
- ☑ Makes it **easy to use** and maintain
- ☑ Provides you **legal protection**
- ☑ **Inter-operates** with other systems using **standards**

Gives ownership and recognition



Henry A. Murray Research Archive at Harvard University

HOME POLICIES CONTACT US

Quick Links

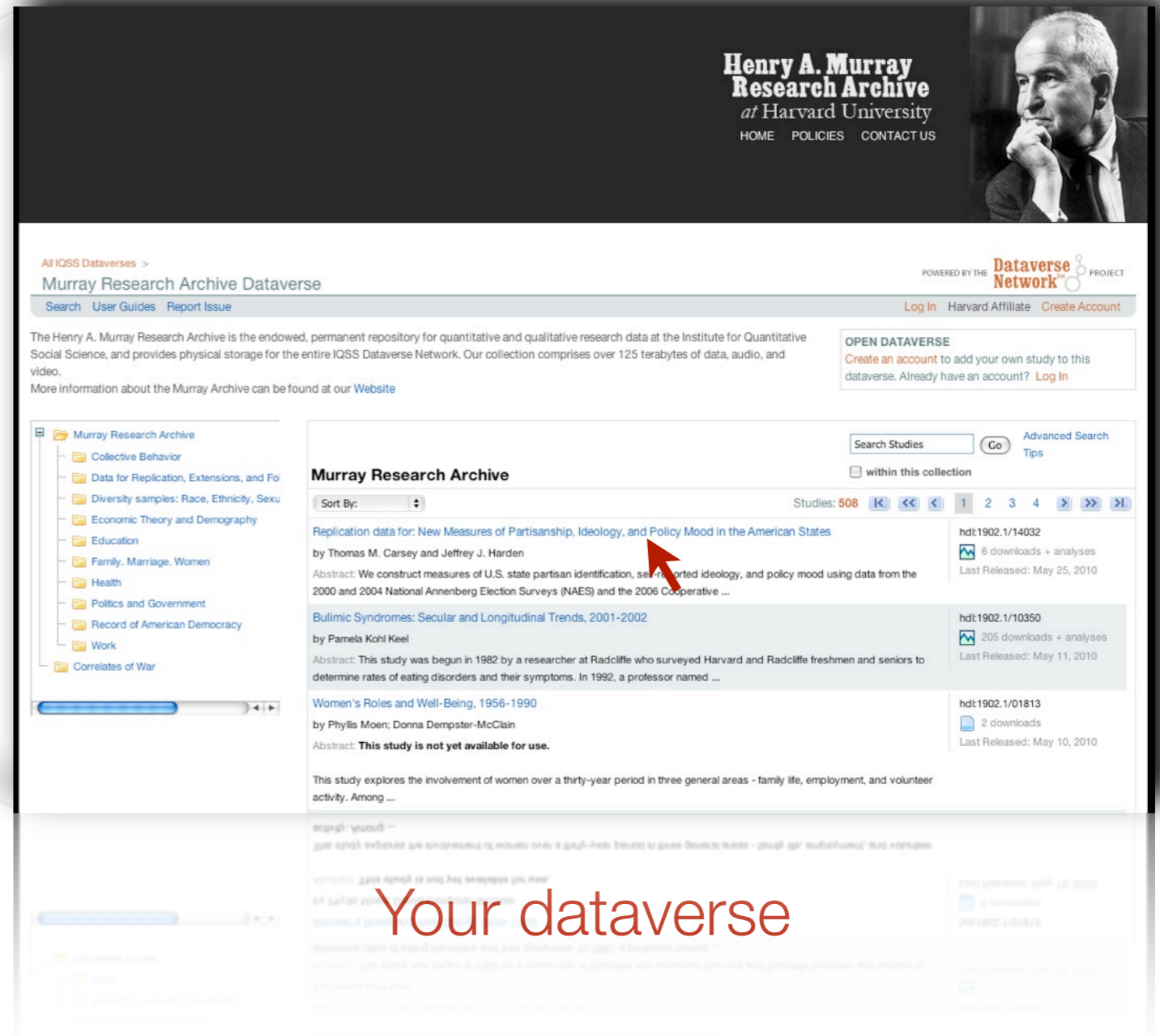
- Find Data
- Deposit Data
- Dissertation Grants

Home

The Henry A. Murray Research Archive is Harvard's 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 100 terabytes of data, audio, and video. We preserve in perpetuity all types of data of interest to the research community, including numerical, video, audio, interview notes, and other data. We accept data deposits through this web site, which is powered by our Dataverse Network software.

The Henry A. Murray Archive, through its endowment, supports permanent bit-level preservation of all social science research studies directly deposited in the IQSS Dataverse Network. (For details, see the Murray Archive Preservation policies.)

Your web site



Henry A. Murray Research Archive at Harvard University

HOME POLICIES CONTACT US

POWERED BY THE Dataverse Network PROJECT

Murray Research Archive Dataverse

Search User Guides Report Issue

Log In Harvard Affiliate Create Account

OPEN DATVERSE
Create an account to add your own study to this dataverse. Already have an account? Log In

Murray Research Archive

Sort By: [dropdown]

Studies: 508

Replication data for: New Measures of Partisanship, Ideology, and Policy Mood in the American States
by Thomas M. Carsey and Jeffrey J. Harden
Abstract: We construct measures of U.S. state partisan identification, self-reported ideology, and policy mood using data from the 2000 and 2004 National Annenberg Election Surveys (NAES) and the 2006 Cooperative ...

Bulimic Syndromes: Secular and Longitudinal Trends, 2001-2002
by Pamela Kohl Keel
Abstract: This study was begun in 1982 by a researcher at Radcliffe who surveyed Harvard and Radcliffe freshmen and seniors to determine rates of eating disorders and their symptoms. In 1992, a professor named ...

Women's Roles and Well-Being, 1956-1990
by Phyllis Moen; Donna Dempster-McClain
Abstract: This study is not yet available for use.

This study explores the involvement of women over a thirty-year period in three general areas - family life, employment, and volunteer activity. Among ...

Your dataverse

Gives ownership and recognition

All IQSS Dataverses > **Murray Research Archive Dataverse** POWERED BY THE **Dataverse Network** PROJECT

[Search](#) [User Guides](#) [Report Issue](#) [Log In](#) [Harvard Affiliate](#) [Create Account](#)

REPLICATION DATA FOR: NEW MEASURES OF PARTISANSHIP, IDEOLOGY, AND POLICY MOOD IN THE AMERICAN STATES

hdl:1902.1/14032 UNF:5:EKgHvTNfkkS86dNzABihNw==
Version: 2 – Released: Tue May 25 10:53:02 EDT 2010

[Cataloging Information](#) [Documentation, Data and Analysis](#) [User Comments](#) [Versions](#)

How to Cite [Why cite?](#)

In order to credit those who developed and distributed the data, select either "print" or "web" citation format, copy the citation generated and paste into your document or website. This citation contains a persistent identifier (and in some cases, a Numerical Fingerprint) for permanent reference. Include the citation wherever the data is referenced.


Please use this citation

```
Thomas M. Carsey and Jeffrey J. Harden, 2010, "Replication data for: New Measures of Partisanship, Ideology, and Policy Mood in the American States",  
http://hdl.handle.net/1902.1/14032 UNF:5:EKgHvTNfkkS86dNzABihNw== Murray Research Archive [Distributor] V2 [Version]
```

Citation to the author **Citation to the distributor**

Citation Format

Citation Information

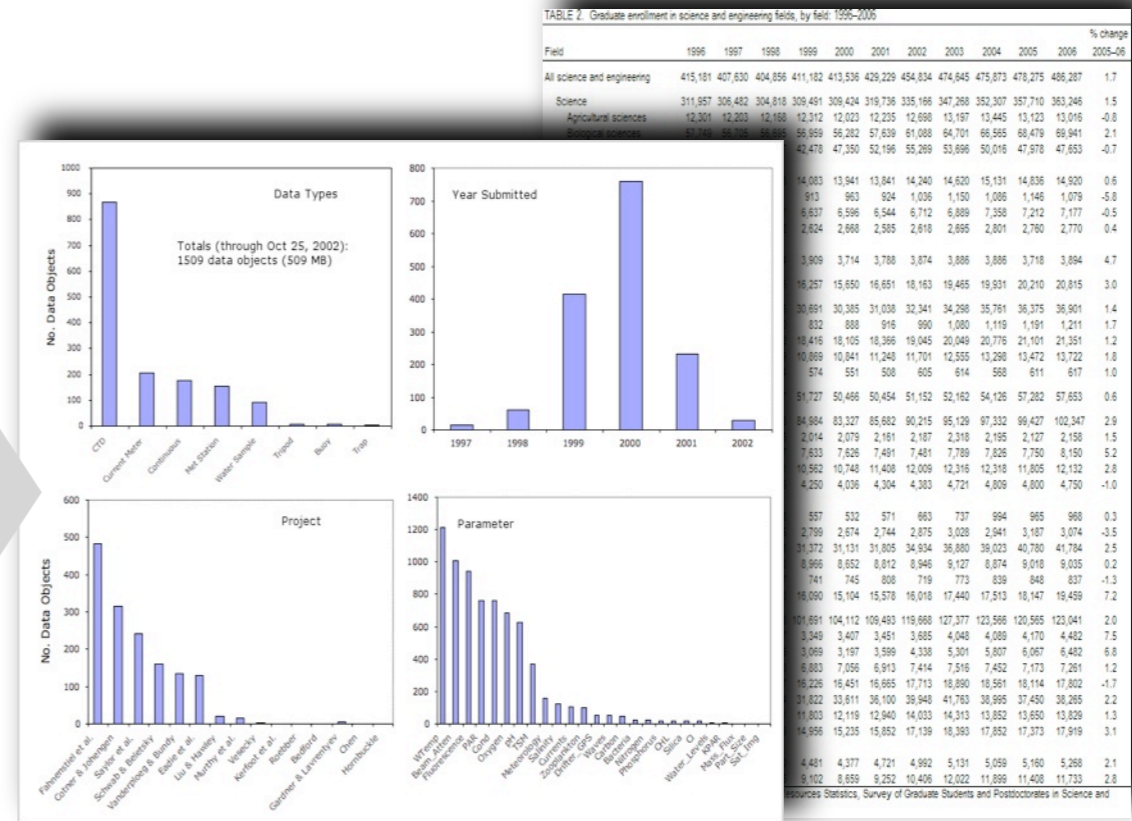
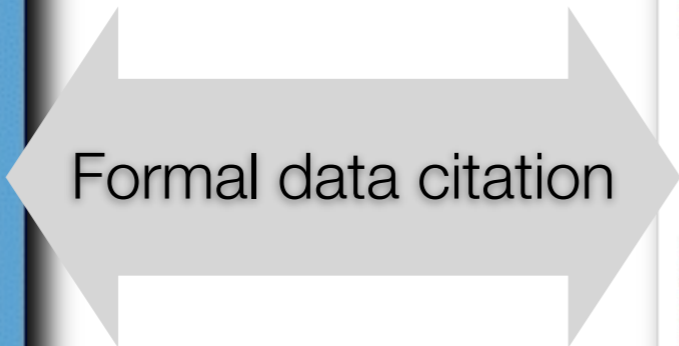
Study Global Id	hdl:1902.1/14032
Authors	Thomas M. Carsey and Jeffrey J. Harden (University of North Carolina at Chapel Hill)
Production Date	2010
Distributor	Murray Research Archive 
Distributor Contact	Jeffrey J. Harden (University of North Carolina at Chapel Hill), jjharden@unc.edu
Deposit Date	January 18, 2010
Replication For	Carsey, Thomas M. and Jeffrey J. Harden. 2010. "New Measures of Partisanship, Ideology, and Policy Mood in the American States." State Politics and Policy Quarterly 10(2):136-156.
Provenance	Murray Research Archive Dataverse

Abstract and Scope

[Abstract](#)

We construct measures of U.S. state partisan identification, self-reported ideology, and policy mood using data from the 2000 and 2004 National Annenberg Election Surveys (NAES) and the 2006 Cooperative Congressional Election Study (CCES). These measures improve on existing methods for estimating state-level preferences because the surveys provide larger state samples without pooling across years. After detailing our methods for constructing the measures, we assess their validity through comparisons with measures already in use by scholars

✓ Persistent Citation for Data



Journal article, book

Dataset, code, documentation to replicate results

Please use this citation

Jacquelyn Boone James; Rosalind C. Barnett, 2003, "The Aspirations and Experiences of Undergraduates, 2000", <http://hdl.handle.net/1902.1/00092> UNF:3:0sjRWertkBtT9wlgcuP4Fg== Murray Research Archive [Distributor] V1 [Version]

Citation Format

A persistent identifier and url which never changes (hdl:1902.2/7768):

<http://handle.net/1902.2/7768>

Universal Numerical Fingerprint to verify data set:

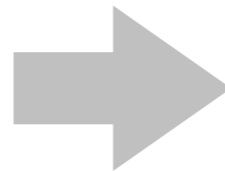
UNF:3:upor0mnBeXa9fujUkPqNNA==

☑ Convert Data to a Preservable and Verifiable Format

Data Set File

Metadata and Format Information

1	4	4	21	...	121
1	2	2	91	...	212
1	9	2	72	...	104
0	2	2	2	...	321
1	6	2	12	...	204
1	9	4	52	...	311
0	3	2	23	...	92
0	2	5	91	...	212
0	5	8	91	...	91
1	9	1	72	...	104
⋮	⋮	⋮	⋮	⋮	⋮
1	2	2	91	...	212



1. Extract variable name, description and summary statistics
2. Convert data set to a preservation format, independent of the software package
3. Apply a cryptographic algorithm to canonical format
4. Get alphanumeric string based on semantic contents of the digital object:
 - uniquely summarizes the contents,
 - but does not convey its information



UNF:5:EKgHvTNfkkS86dNzABlhNw==

Change content



Changes UNF

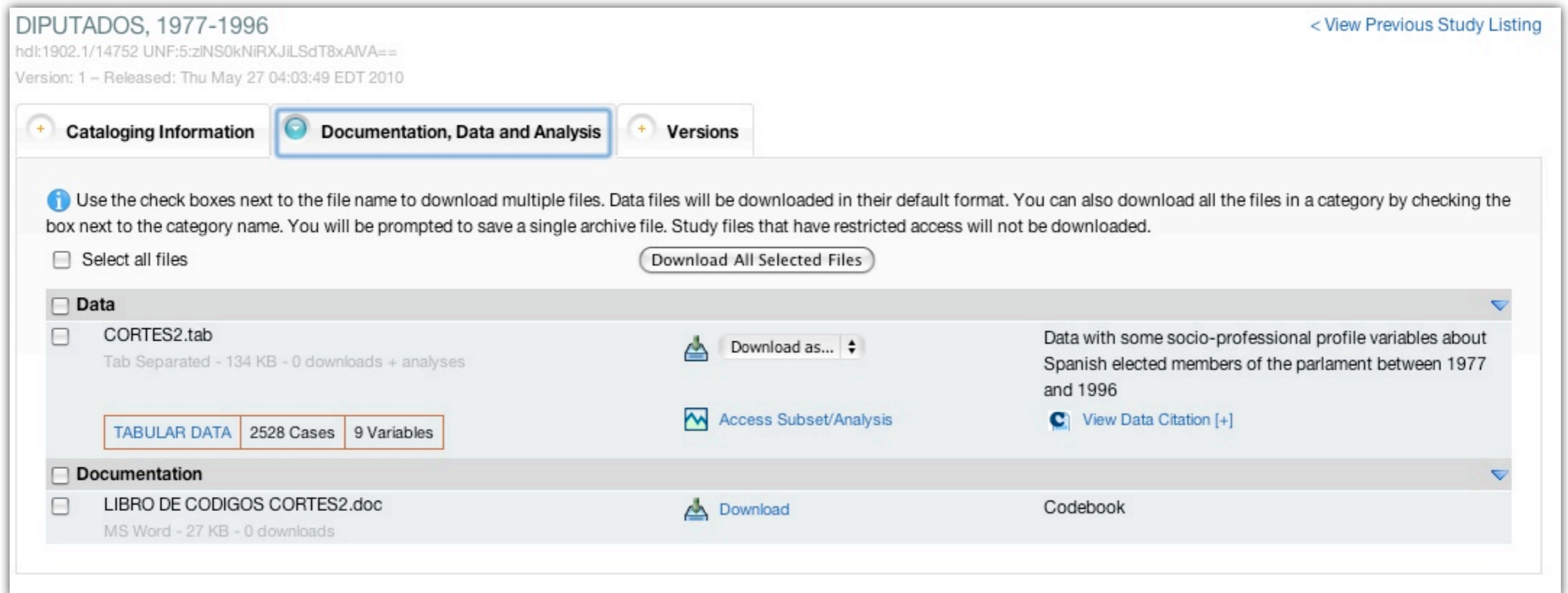
Change format



Doesn't change UNF

Distributes to the public, but also supports restricted access

Studies and files can be completely public ...



The screenshot shows a web interface for a study titled "DIPUTADOS, 1977-1996". The interface includes a header with the study title, a "View Previous Study Listing" link, and three tabs: "Cataloging Information", "Documentation, Data and Analysis" (which is selected), and "Versions". Below the tabs is an information box with a checkmark icon and text explaining that users can use checkboxes to download multiple files, but restricted access files will not be downloaded. A "Download All Selected Files" button is present. The main content area is divided into two sections: "Data" and "Documentation". The "Data" section contains a file named "CORTES2.tab" (134 KB) with a "Download as..." button and a description: "Data with some socio-professional profile variables about Spanish elected members of the parliament between 1977 and 1996". It also features a "View Data Citation [+]" link and a summary box showing "TABULAR DATA", "2528 Cases", and "9 Variables". The "Documentation" section contains a file named "LIBRO DE CODIGOS CORTES2.doc" (27 KB) with a "Download" button and a description: "Codebook".

Everybody can explore the cataloging information (metadata), download data and documentation files, and subset and analyze data if available.

Distributes to the public, but also supports restricted access

... or some files can be restricted ...

THE ASPIRATIONS AND EXPERIENCES OF UNDERGRADUATES, 2000 [< View Previous Study Listing](#)











hdl:1902.1/00092 UNF:3:0sjRWerlkBtT9wlgcuP4Fg==
Version: 1 - Released: Mon May 10 13:02:44 EDT 2010

[+ 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.

! Access to some files is restricted, and those files are not available for downloading. Check the [Terms of Use](#) for more information.

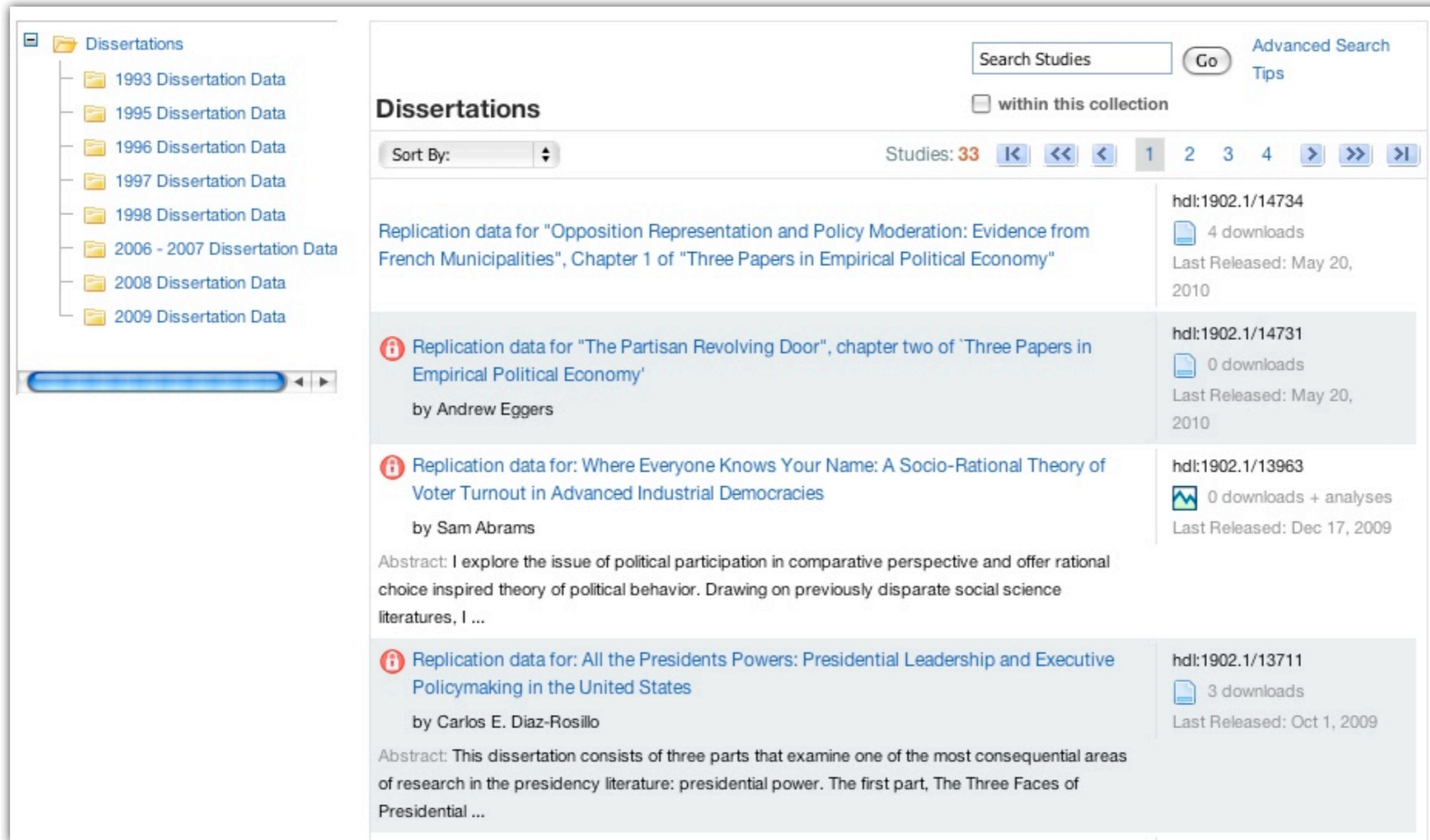
Select all files [Download All Selected Files](#)

<input type="checkbox"/> 1. Documentation			
<input type="checkbox"/> 00092JamesBarnett.zip ZIP Archive - 562 KB - 18 downloads	 Download		Codebooks and appendices for this study
<input type="checkbox"/> 00092James-Undergraduates-Codebooks.pdf Adobe PDF - 562 KB - 25 downloads	 Download		Description of coded data variables
<input type="checkbox"/> 00092James-Undergraduates-Measures.pdf Adobe PDF - 1 MB - 2 downloads	 Download		Collection of blank measures used in the study
<input type="checkbox"/> 2. Data			
<input type="checkbox"/> 00092James-Undergraduates-first_yr2.tab Tab Separated - 180 KB - 4 downloads + analyses TABULAR DATA 480 Cases 179 Variables	 Restricted Access  Access Subset/Analysis	 View Data Citation [+]	Data file - first year students
<input type="checkbox"/> 00092James-Undergraduates-seniors2.tab Tab Separated - 143 KB - 3 downloads + analyses TABULAR DATA 325 Cases 200 Variables	 Restricted Access  Access Subset/Analysis	 View Data Citation [+]	Data file - seniors
<input type="checkbox"/> Text Files.zip ZIP Archive - 401 KB - 3 downloads	 Restricted Access		Electronic text files of TATs

- ▶ Only authorized users or groups can access the files
- ▶ Groups can be based on IP address

Distributes to the public, but also supports restricted access

... or the entire study can be restricted.



The screenshot shows a digital library interface. On the left is a sidebar with a tree view of folders labeled 'Dissertations', containing sub-folders for years from 1993 to 2009. The main content area is titled 'Dissertations' and includes a search bar with the text 'Search Studies', a 'Go' button, and a link to 'Advanced Search Tips'. Below the search bar is a checkbox for 'within this collection'. A navigation bar shows 'Sort By:' with a dropdown arrow, 'Studies: 33', and page navigation buttons (1, 2, 3, 4, >, >>, >>>). The list of studies includes:

- Replication data for "Opposition Representation and Policy Moderation: Evidence from French Municipalities", Chapter 1 of "Three Papers in Empirical Political Economy" (hdl:1902.1/14734, 4 downloads, Last Released: May 20, 2010)
- Replication data for "The Partisan Revolving Door", chapter two of 'Three Papers in Empirical Political Economy' by Andrew Eggers (hdl:1902.1/14731, 0 downloads, Last Released: May 20, 2010)
- Replication data for: Where Everyone Knows Your Name: A Socio-Rational Theory of Voter Turnout in Advanced Industrial Democracies by Sam Abrams (hdl:1902.1/13963, 0 downloads + analyses, Last Released: Dec 17, 2009)
- Replication data for: All the Presidents Powers: Presidential Leadership and Executive Policymaking in the United States by Carlos E. Diaz-Rosillo (hdl:1902.1/13711, 3 downloads, Last Released: Oct 1, 2009)

Restricted studies can be searched but user cannot access the study page unless has the right permissions

Indexes all metadata for quick data discovery

Quick and advanced search of all metadata fields:

- Title
- Authors
- Distributors
- Abstract
- keywords
- Years
- Geographic information
- Data collection
- Terms of use
- ...
- Variable labels
- Variable descriptions

Uses **Lucene** index server as the search engine.

✔ Supports subsetting and analysis for (some) data files

REPLICATION DATA FOR: IMPROVING ANCHORING VIGNETTES: DESIGNING SURVEYS TO CORRECT INTERPERSONAL INCOMPARABILITY [Back to Study](#)

DATA FILE: HOPKINSDATANOZIP.TAB

Download Subset **Recode & Case-Subset** **Descriptive Statistics** **Advanced Statistical Analysis**

Selected Variables

- ZIP
- CASEID
- SPEED1
- SPEED2

For SPSS and STATA Files

Select variables from table below (select)

Variable Information Table

<input type="checkbox"/>	Type		
<input checked="" type="checkbox"/>	Discrete		
<input checked="" type="checkbox"/>	Discrete		
<input type="checkbox"/>	Discrete	HOPRAND1	HOPRAND1
<input type="checkbox"/>	Discrete	MUCHSAY1	MUCHSAY1

Choose a Statistical Model

- Categorical Data Analysis**
 - Cross-Tabulation**
- Ecological Inference Models**
 - Hierarchical Multinomial-Dirichlet Ecological Inference Model for R x C Tables
- Event Count Models**
 - Negative Binomial Reg for Event Count Dep Vars
 - Poisson Reg for Event Count Dep Vars
- Models for Continuous Bounded Dependent Variables**
 - Exponential Reg for Duration Dep Vars
 - Gamma Reg for Cont, Positive Dep Vars
 - Log-Normal Reg for Duration Dep Vars
 - Weibull Reg for Duration Dep Vars
- Models for Continuous Dependent Variables**
 - Least Squares Reg for Cont Dep Vars
 - Linear regression for Left-Censored Dep Variable
- Models for Dichotomous Dependent Variables**
 - Logistic Reg for Binary Dep Vars
 - Probit Reg for Binary Dep Vars
 - Rare Events Logistic Reg for Binary Dep Vars
- Models for Ordinal Dependent Variables**
 - Ordinal Logistic Reg for Ordered Cat Dep Vars
 - Ordinal Probit Reg for Ordered Cat Dep Vars

Show
10 Variables

1 2 3 4 > >> >|

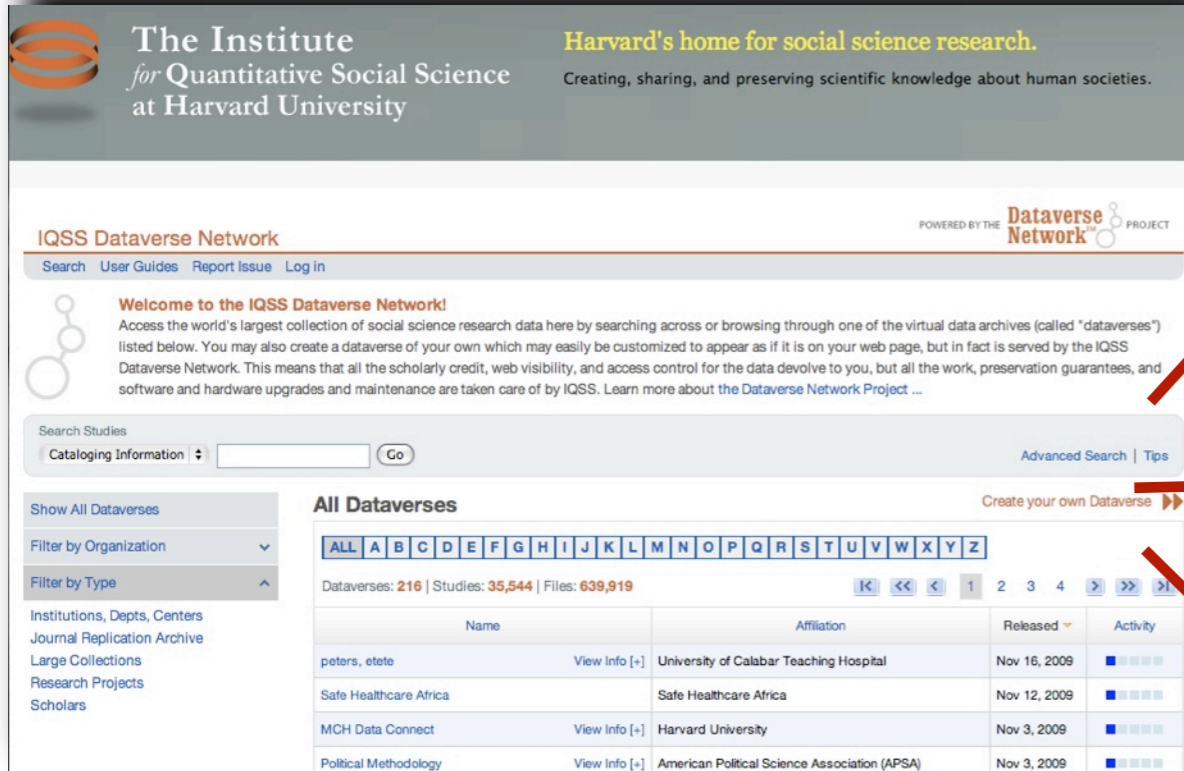
Summary

-
-
-
-

 Makes it easy to use and maintain

A dataverse = a site for publishing your data

The Dataverse Network = the software



The Institute for Quantitative Social Science at Harvard University. Harvard's home for social science research. Creating, sharing, and preserving scientific knowledge about human societies.

POWERED BY THE **Dataverse Network** PROJECT

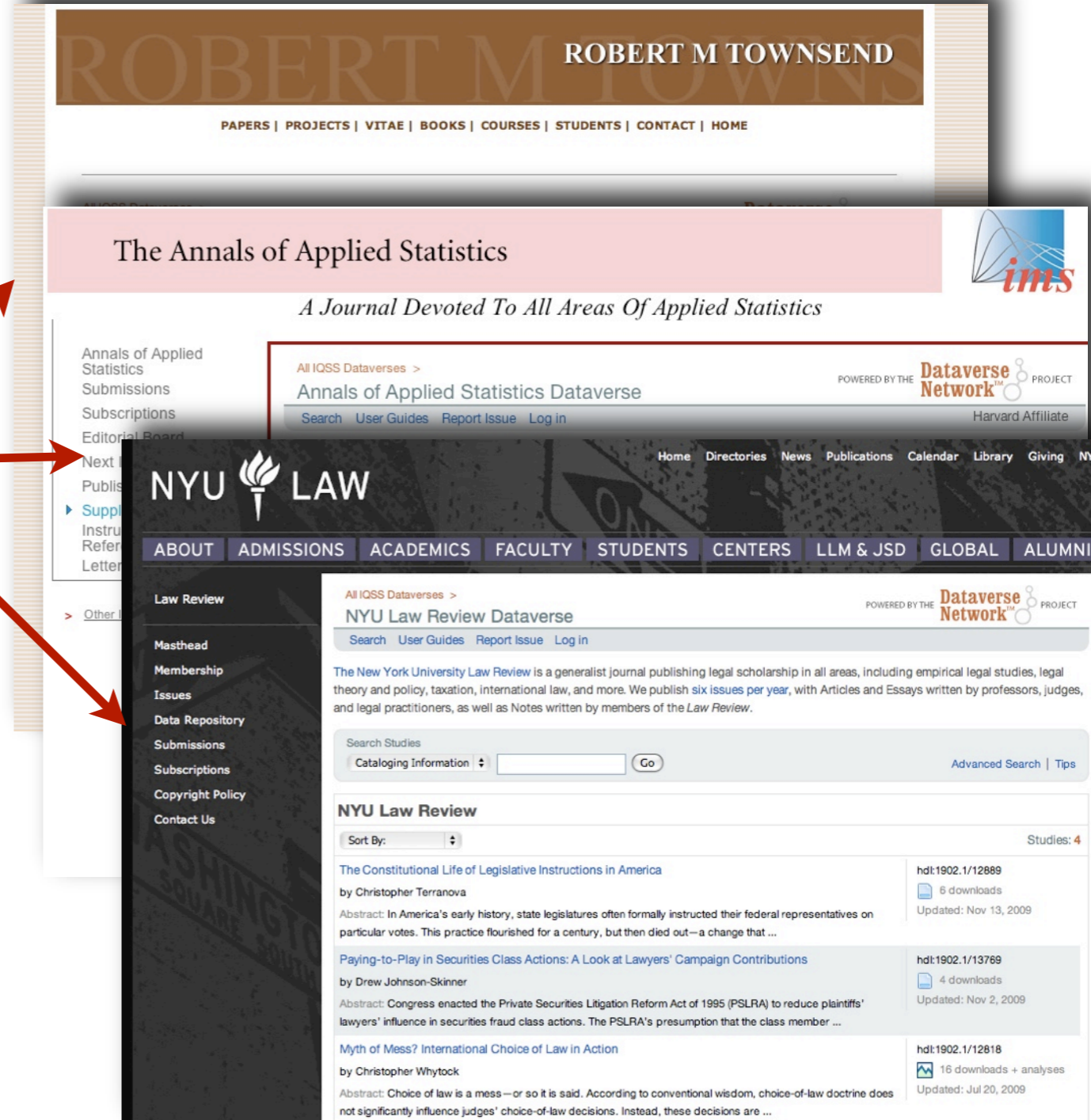
Welcome to the IQSS Dataverse Network! Access the world's largest collection of social science research data here by searching across or browsing through one of the virtual data archives (called "dataverses") listed below. You may also create a dataverse of your own which may easily be customized to appear as if it is on your web page, but in fact is served by the IQSS Dataverse Network. This means that all the scholarly credit, web visibility, and access control for the data devolve to you, but all the work, preservation guarantees, and software and hardware upgrades and maintenance are taken care of by IQSS. Learn more about the Dataverse Network Project ...

Search Studies
Cataloging Information Go Advanced Search | Tips

All Dataverses Create your own Dataverse

Dataverses: 216 | Studies: 35,544 | Files: 639,919

Name	Affiliation	Released	Activity
peters, etete View Info [+]	University of Calabar Teaching Hospital	Nov 16, 2009	■ ■ ■ ■ ■
Safe Healthcare Africa	Safe Healthcare Africa	Nov 12, 2009	■ ■ ■ ■ ■
MCH Data Connect View Info [+]	Harvard University	Nov 3, 2009	■ ■ ■ ■ ■
Political Methodology View Info [+]	American Political Science Association (APSA)	Nov 3, 2009	■ ■ ■ ■ ■



ROBERT M TOWNSEND
PAPERS | PROJECTS | VITAE | BOOKS | COURSES | STUDENTS | CONTACT | HOME

The Annals of Applied Statistics
A Journal Devoted To All Areas Of Applied Statistics

POWERED BY THE **Dataverse Network** PROJECT
Harvard Affiliate

NYU LAW
Home Directories News Publications Calendar Library Giving NYU

ABOUT | ADMISSIONS | ACADEMICS | FACULTY | STUDENTS | CENTERS | LLM & JSD | GLOBAL | ALUMNI

NYU Law Review
The New York University Law Review is a generalist journal publishing legal scholarship in all areas, including empirical legal studies, legal theory and policy, taxation, international law, and more. We publish six issues per year, with Articles and Essays written by professors, judges, and legal practitioners, as well as Notes written by members of the Law Review.

Search Studies
Cataloging Information Go Advanced Search | Tips

NYU Law Review Studies: 4

Sort By:

The Constitutional Life of Legislative Instructions in America by Christopher Terranova Abstract: In America's early history, state legislatures often formally instructed their federal representatives on particular votes. This practice flourished for a century, but then died out—a change that ...	hdt:1902.1/12889 6 downloads Updated: Nov 13, 2009
Paying-to-Play in Securities Class Actions: A Look at Lawyers' Campaign Contributions by Drew Johnson-Skinner Abstract: Congress enacted the Private Securities Litigation Reform Act of 1995 (PSLRA) to reduce plaintiffs' lawyers' influence in securities fraud class actions. The PSLRA's presumption that the class member ...	hdt:1902.1/13769 4 downloads Updated: Nov 2, 2009
Myth of Mess? International Choice of Law in Action by Christopher Whytock Abstract: Choice of law is a mess—or so it is said. According to conventional wisdom, choice-of-law doctrine does not significantly influence judges' choice-of-law decisions. Instead, these decisions are ...	hdt:1902.1/12818 16 downloads + analyses Updated: Jul 20, 2009

- A Dataverse Network holds multiple dataverses
- Each dataverse can be created in two clicks and customized as you want
- Your dataverse may include your data or links to data from other dataverses

Provides you Legal Protection

TERMS OF USE

DATAVERSE NETWORK STUDY CREATION TERMS OF USE

The following terms and conditions govern all use of the IQSS Dataverse website (<http://dvn.iq.harvard.edu>) (**the Site**) and the services available on or at the Site (taken together, **the Service**). The Site is owned and operated by IQSS, of Harvard University (**IQSS**). The Service provides you with the ability to post User Submissions (**the "Materials"**) to virtual archives (**the "Archive"**) administered by third parties (**the "Archive Administrators"**) subject to the permissions and terms of use required by that Archive.

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

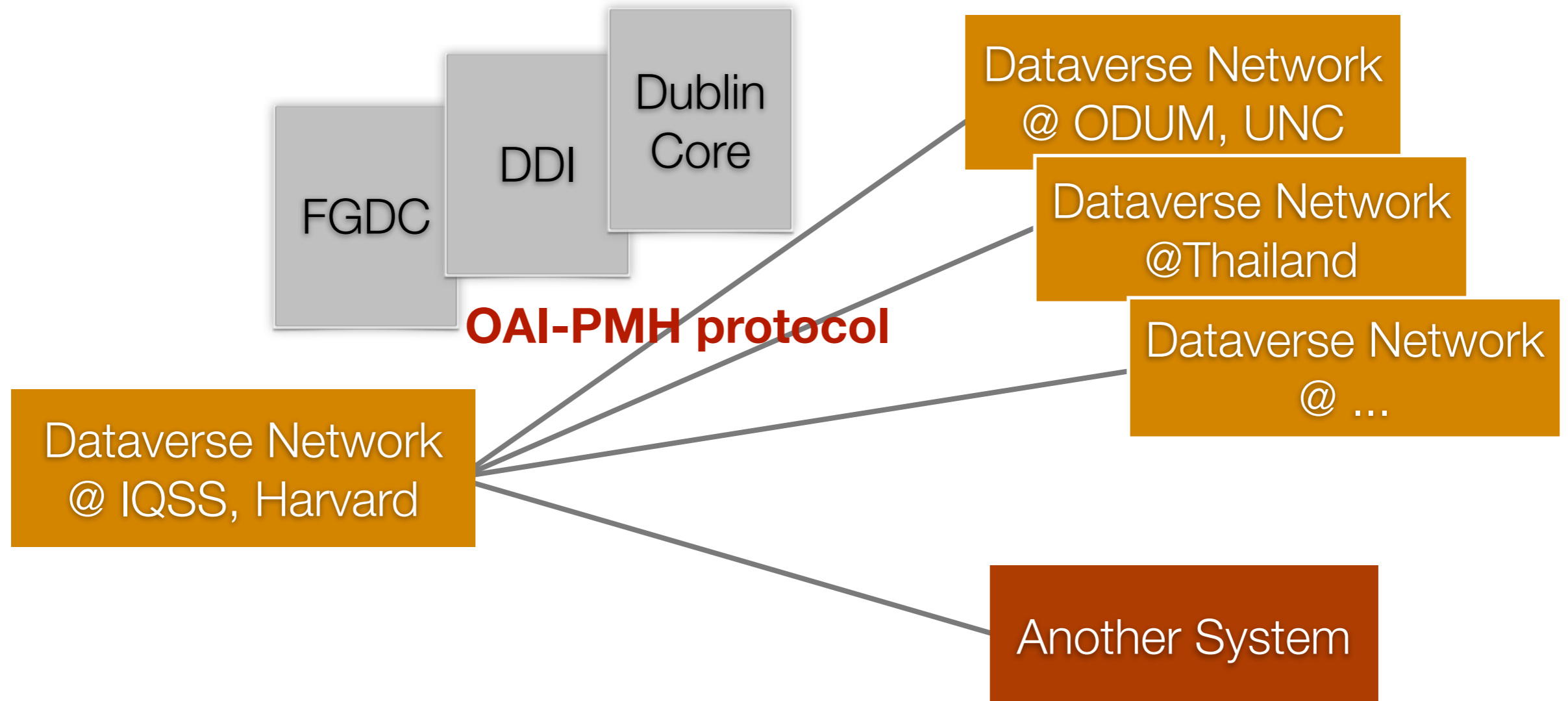
I agree and accept these terms of use.

Agreement for Data Deposit

- ...
- ✓ If human subjects were studied in the collection of the Materials, you collected the Materials with IRB approval;
 - ✓ The Materials do not contain high-risk confidential information such as social security numbers; credit card numbers; medical record numbers; health plan numbers; other account numbers; certificate/license numbers; or biometric identifiers (fingerprints, retina, voice print, DNA etc.).
- ...

- The terms of Use in the Harvard Dataverse Network are vetted by lawyers from the Harvard's Office of the General Counsel
- You can customize the terms of use data deposit and data downloads in your dataverse or entire Dataverse Network

Inter-operates with other systems



- Study and variable metadata is exported into XML (Dublin Core, Data Descriptive Initiative - DDI, and others) and MARC.
- Supports OAI-PMH, z39.50, LOCKKS
- Plan to add RDF support in the future

New Functionality

- **Versioning:** Never lose previous versions of studies
- **Deaccessioning:** Never lose track of previous holding
- **Open access options:** Enable contributors to collaboratively edit
- **User comments:** Collect user comments about data
- **UNF 5.0:** Advanced semantic fingerprints for your data
- **Social Network Data:** Deposit, format, analyze, and preserve social networks
- **LOCKSS compatibility:** Automatic replication for preservation

Thanks!

Next...

We invite you to participate in a workshop on **Data Publication and Citation Principles** at IQSS, Harvard University

Organizers: Micah Altman and Mercè Crosas

Contact me at mcrosas@hmdc.harvard.edu if you are interested.

<http://thedata.org> (The project site)

<http://dvn.iq.harvard.edu> (The Dataverse Network at Harvard)

Acknowledgements:

Gary King, Micah Altman, Gustavo Durand, Ellen Kraffmiller, Kevin Condon, Bob Treacy, Leonid Andrev, Michael Heppler, Steve Kraffmiller