

The User Experience in a Dataverse Supported Repository

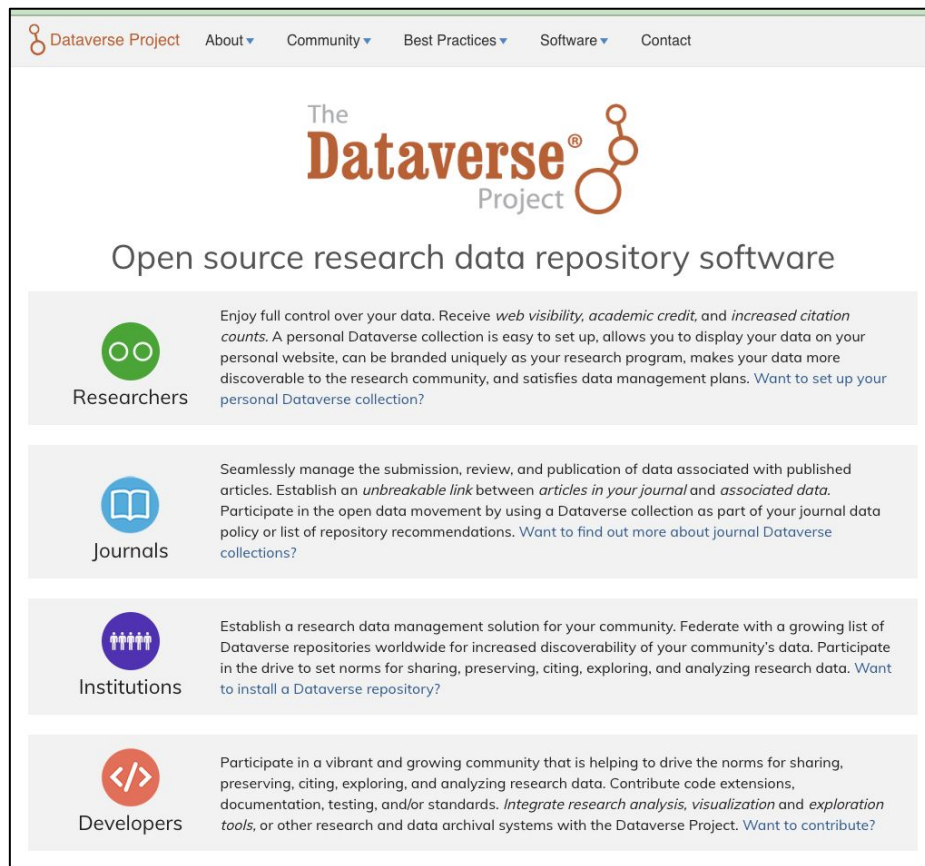
Sonia Barbosa

National Yang Ming Chiao Tung University, February 2024

Agenda

AGENDA: The User Experience

- Community
- Collections
- Datasets
- Files
- Files and External Tools
- Permissions and Access
- Data Exploration
- Search and Browse Functionality



The screenshot shows the homepage of the Dataverse Project. At the top is a navigation bar with the logo and links for 'About', 'Community', 'Best Practices', 'Software', and 'Contact'. The main heading reads 'The Dataverse Project' with a logo of three connected circles. Below this is the tagline 'Open source research data repository software'. The page features four main sections, each with an icon and a brief description:

- Researchers:** Represented by a green circle icon with two white dots. Text: 'Enjoy full control over your data. Receive *web visibility, academic credit, and increased citation counts*. A personal Dataverse collection is easy to set up, allows you to display your data on your personal website, can be branded uniquely as your research program, makes your data more discoverable to the research community, and satisfies data management plans. [Want to set up your personal Dataverse collection?](#)
- Journals:** Represented by a blue book icon. Text: 'Seamlessly manage the submission, review, and publication of data associated with published articles. Establish an *unbreakable link* between *articles in your journal* and *associated data*. Participate in the open data movement by using a Dataverse collection as part of your journal data policy or list of repository recommendations. [Want to find out more about journal Dataverse collections?](#)
- Institutions:** Represented by a purple circle icon with four white figures. Text: 'Establish a research data management solution for your community. Federate with a growing list of Dataverse repositories worldwide for increased discoverability of your community's data. Participate in the drive to set norms for sharing, preserving, citing, exploring, and analyzing research data. [Want to install a Dataverse repository?](#)
- Developers:** Represented by a red circle icon with white code symbols. Text: 'Participate in a vibrant and growing community that is helping to drive the norms for sharing, preserving, citing, exploring, and analyzing research data. Contribute code extensions, documentation, testing, and/or standards. *Integrate research analysis, visualization and exploration tools*, or other research and data archival systems with the Dataverse Project. [Want to contribute?](#)

Dataverse Repository Features

- Support for FAIR Data Principles
- Data citation for datasets and files
- OAI-PMH (Harvesting)
- APIs for interoperability and custom integrations
- API client libraries
- DataCite integration
- Login via Shibboleth
- Login via ORCID, Google, GitHub, or Microsoft
- Login via OpenID Connect (OIDC)
- Internationalization
- Versioning
- Restricted files
- Embargo
- Custom licenses
- Custom terms of use
- Publishing workflow support
- File hierarchy
- File previews
- Preview and analysis of tabular files
- Usage statistics and metrics
- Guestbook
- Fixity checks for files
- File download in R and TSV format
- Faceted search
- Customization of collections
- Private URL
- Widgets
- Notifications
- Schema.org JSON-LD
- External tools
- External vocabulary
- Dropbox integration
- GitHub integration
- Integration with Jupyter notebooks
- User management
- Curation status labels
- Branding
- Backend storage on S3 or Swift
- Direct upload and download for S3
- Export data in BagIt format
- Post-publication automation (workflows)
- Pull header metadata from Astronomy (FITS) files
- Provenance
- Support for rsync
- Auxiliary files for data files

<https://dataverse.org/software-features>

The Dataverse Project and FAIR

FAIR Principles and Dataverse

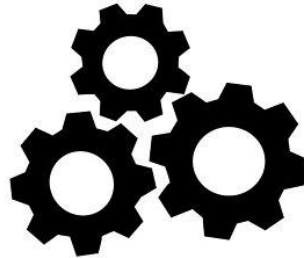
F
Findable



A
Accessible



I
Interoperable



R
Reusable



https://en.wikipedia.org/wiki/File:FAIR_data_principles.jpg

How Dataverse Supports FAIR

- **FINDABLE:** Persistent Identifiers (DOIs, Handles) for datasets (and files), and enabling metadata indexing by search engines
- **ACCESSIBLE:** Open access to (meta)data and ensures data can be downloaded in machine-readable formats
- **INTEROPERABLE:** Provides standardised metadata schemas and enabling the integration of data with research tools and platforms
- **REUSABLE:** Licenses that clearly state how data can be used and ensuring that data are well documented and preserved for long term use

The User Experience - an Overview

IQSS / dataverse

Code Issues 1.1k Pull requests 106 Actions Projects Security 26 Insights Settings

dataverse Public

develop 256 Branches 67 Tags

Go to file

pdurbin Merge pull request #10250 from IQSS/10249_search_api_typo

- .github Add default label to request issue to
- conf Merge branch 'develop' into 6783-s
- doc Merge pull request #10250 from IQSS/10249_search_api_typo
- modules Bump version to 6.1
- scripts Add configuration for automatic XH
- src Merge pull request #10269 from O
- tests add S3AccessIT to the list of integr
- env update Solr 8.11.1 to 9.3.0 (minus c
- .gitattributes Add the .gitattributes file to set LF
- .gitignore remove Vagrant (unused) #9838
- .readthedocs.yml ci(docs): make RTFD install graphviz, too

docker-compose-dev.yml Update docker-compose-dev.yml 4 days ago

pom.xml Merge remote-tracking branch 'origin/develop' into 9920... last month

README Code of conduct License Security

Groups

Conversations Search conversations within dataverse-com...

Dataverse Users Community 982 members

Welcome to the Dataverse Users Community Group! Please feel free to ask a question, share feedback or start a discussion about anything and everything involving Dataverse. For more information about the Dataverse Project please visit: <https://dataverse.org>

shla...@gmail.com, Philipp Konzett 4 Looking for Dataverse - OSF Addon Users to Interview - Hi all, I've scheduled a meeting with OSF, too. At UI/DataverseNO we started a pilot on OSF 2-3 Jan 27 ☆

yuej...@gmail.com, ...shla...@gmail.com 4 Removed DOI display - Your example shows a file DOI. For one thing, you may want to reconsider issuing DOIs for files. UVA Jan 26 ☆

courtne...@gmail.com, Julian Gautier 2 DataCite updates and Dataverse - Hi Courtney, I agree, discussing during a community call would be helpful. Specifically I've been Jan 26 ☆

yuej...@gmail.com, Jim Myers 7 curl command to change - Hi Jim, Thank you so much for the detail guide. We will try this way. Best, Yue On Friday, January 26 Jan 26 ☆

Philip Durbin top 25 cited datasets in your Dataverse installation - I don't know if this is common knowledge or not, but I was just chatting with Kelly Statis from Jan 26 ☆

naima...@gmail.com, don.si...@gmail.com

victor...@gmail.com, Jim Myers 4

cs...@pclinic.cl, Philip Durbin 2

cs...@pclinic.cl, Philip Durbin 2

dsae...@gmail.com, ... Philip Durbin 6

Matthew Mangun, cs...@pclinic.cl 2

Report repository

Releases 66

+170 contributors

Deployments 32

beta-testing

+31 deployments

Languages

- Java 82.6%
- JavaScript 3.3%
- Python 0.8%
- Other 1.2%
- HTML 10.2%
- Shell 1.3%
- XSLT 0.6%

Dataverse®

Dataverse is an [open source](#) software platform for sharing, finding, citing, and preserving research data (developed by the [Dataverse team](#) at the [Institute for Quantitative Social Science](#) and the [Dataverse community](#)).

[dataverse.org](#) is our home on the web and shows a map of Dataverse installations around the world, a list of [features](#), [integrations](#) that have been made possible through [REST APIs](#), our development [roadmap](#), and more.

We maintain a demo site at [demo.dataverse.org](#) which you are welcome to use for testing and evaluating Dataverse.

To install Dataverse, please see our [Installation Guide](#) which will prompt you to download our [latest release](#).

To discuss Dataverse with the community, please join our [mailing list](#), participate in a [community call](#), chat with us at [chat.dataverse.org](#), or attend our annual [Dataverse Community Meeting](#).

We love contributors! Please see our [Contributing Guide](#) for ways you can help.

Dataverse is a trademark of President and Fellows of Harvard College and is registered in the United States.

[API Test Status](#) [API Test Coverage](#) [inaccessible](#) [Maven Tests](#) [passing](#) [Unit Test Coverage](#) [20%](#) [Guides Build Status](#) [passing](#)

Home

Explore

Notifications

Messages

Grok

Lists

Bookmarks

Communities

Premium

Profile

More

Post

For you Following

What is happening?!

Post

Slava Tykhonov @4tykxhonov · Jan 27 · 🌐

My latest experiments on the Open Source #LLM models "decompression" show great potential transforming all "packaged" content into a structured format and ingesting it back into the knowledge graph. Even in the current state of art, Generative AI is capable of saving...

Skosmos

Queen Beatrix Thesaurus

Made in Heaven

Slava Tykhonov @4tykxhonov · 1h

It looks like this can be considered as official start of the battle between human and #AI in the field of information classification and ontologies. And we're in the very beginning...

RRE

Infectious disease

Who to follow

OpenCitations @opencitations Follow

CORE @oacore Follow

Melissa Sands @melissaless Follows you Follow

Terms of Service Privacy Policy Cookie Policy Accessibility Ads Info More ... © 2024 X Corp.

➔ Log In

Log in or sign up with your institutional account — more [information about account creation](#). Leaving your institution? Please contact [Harvard Dataverse Support](#) for assistance.

Your Institution



Harvard University

Please select...



Continue

[Allow me to type the name of my institution](#)

Other options

Username/Email

GitHub

Google

ORCID

[Sign up for a Dataverse account.](#)

Collection-level

Current Features



Dataverse Collections

- Own administration
- Own branding (and can be embedded anywhere)

Datasets

- Citation
- Metadata
- Versioning
- Private URL
- Custom Terms/Multiple License/Permissions
- Guestbooks
- Publishing Workflows



Files

- Citation
- Ingest
- Preview/Explore
- Metadata
- Versioning
- Permissions/Embargo/Restrictions



Brain Genomics Superstruct Project (GSP) Dataverse Home
(Harvard University)

Harvard Dataverse >

Contact Share Link Edit

Large scale imaging data sets are necessary to address complex questions regarding the relationship between brain and behavior. The Brain Genomics Superstruct Project Open Access Data Release exposes a carefully vetted collection of neuroimaging, behavior, cognitive, and personality data for over 1,500 human participants. Each neuroimaging data set includes one high-resolution Magnetic Resonance Imaging (MRI) acquisition and one or more resting-state functional MRI acquisitions. Each functional acquisition is accompanied by a fully-automated quality assessment and pre-computed brain morphometrics are also provided.

Search this dataverse... Advanced Search Add Data

Dataverses (0)
Datasets (1)
Files (10)

1 to 1 of 1 Result Sort

Brain Genomics Superstruct Project (GSP)
May 15, 2015

Publication Year
2014 (1)

Subject
Other (1)

Author Name
Buckner, Randy L. (1)
Roffman, Joshua L. (1)
Smoller, Jordan W. (1)

Author Affiliation
Harvard University (1)

Brain Genomics Superstruct Project (GSP)
Buckner, Randy L.; Roffman, Joshua L.; Smoller, Jordan W., 2014, "Brain Genomics Superstruct Project (GSP)",
<https://doi.org/10.7910/DVN/25833>, Harvard Dataverse, V10

Large scale imaging data sets are necessary to address complex questions regarding the relationship between brain and behavior. The Brain Genomics Superstruct Project Open Access Data Release exposes a carefully vetted collection of neuroimaging, behavior, cognitive, and personal...



Brain Genomics Superstruct Project (GSP)

Dataverse (Harvard University)
GSP

General Information

Theme + Widgets

Permissions

Groups

Dataset Templates

Dataset Guestbooks

Featured Dataverses

Dataset-level

Current Features



Dataverse Collections

- Own administration
- Own branding (and can be embedded anywhere)



Datasets

- Citation
- Metadata
- Versioning
- Private URL
- Custom Terms/Multiple License/Permissions
- Guestbooks
- Publishing Workflows



Files

- Citation
- Ingest
- Preview/Explore
- Metadata
- Versioning
- Permissions/Embargo/Restrictions

HARVARD Dataverse

Brain Genomics Superstruct Project (GSP) Dataset

Brain Genomics Superstruct Project (GSP) Dataset (Harvard University)

Harvard Dataverse > Brain Genomics Superstruct Project (GSP) Dataset >

Brain Genomics Superstruct Project (GSP)

Version 10.5

Buckner, Randy L; Roffman, Joshua L; Smoller, Jordan W., 2014, "Brain Genomics Superstruct Project (GSP)", <https://doi.org/10.7910/DVN/25833>, Harvard Dataverse, V10

Cite Dataset Learn about Data Citation Standards.

Access Dataset -
Contact Owner Share

Dataset Metrics
13,076 Downloads

Dataset Metrics

13,076 Downloads

Supported Metadata

Detailed below are what metadata schemas we support for Citation and Domain Specific Metadata in the Dataverse Project:

- **Citation Metadata** (see .tsv version): compliant with DDI Lite, DDI 2.5 Codebook, DataCite 3.1, and Dublin Core's *DCMI Metadata Terms*. Language field uses ISO 639-1 controlled vocabulary.
- **Geospatial Metadata** (see .tsv version): compliant with DDI Lite, DDI 2.5 Codebook, DataCite, and Dublin Core. Country / Nation field uses ISO 3166-1 controlled vocabulary.
- **Social Science & Humanities Metadata** (see .tsv version): compliant with DDI Lite, DDI 2.5 Codebook, and Dublin Core.
- **Astronomy and Astrophysics Metadata** (see .tsv version): These metadata elements can be mapped/exported to the International Virtual Observatory Alliance's (IVOA) VOResource Schema format and is based on Virtual Observatory (VO) Discovery and Provenance Metadata.
- **Life Sciences Metadata** (see .tsv version): based on ISA-Tab Specification, along with controlled vocabulary from subsets of the OBI Ontology and the NCBI Taxonomy for Organisms.
- **Journal Metadata** (see .tsv version): based on the Journal Archiving and Interchange Tag Set, version 1.2.

Subject ID Other
License/Data Use Agreement Custom Dataset Terms

Files Metadata Terms Versions

Search this dataset...

File Type: All Access: All File Tab: All

1 to 10 of 10 Files Download Request Access

GSP_Dataverse_Terms_140422.pdf
Author: GSP, 2014
Published: Aug 26, 2014
1,150 Downloads
Download

GSP_Isa_140500.csv
File Type: CSV+TXT
Published: Aug 26, 2014
194 Downloads
Download

GSP_part10_140500.tar
GSP Version: 10.026
Published: May 21, 2014
1,020 Downloads
Download

Cite Dataset Learn about Data Citation Standards.

Author *
Name *
Barbosa, Sonia
Affiliation *
Harvard University

Identifier Scheme *
Select...
ORCID
ISNI
LCNA
VIAF
GND
DAI
ResearcherID
ScopusID

Contact *
Name *
Affiliation *
Harvard University

Buckner, Randy L.; Roffman, Joshua L.; Smoller, Jordan W., 2014, "Brain Genomics Superstruct Project (GSP)", <https://doi.org/10.7910/DVN/25833>, Harvard Dataverse, V10

Cite Dataset

Learn about Data Citation Standards.

File-level

Current Features



Dataverse Collections

- Own administration
- Own branding (and can be embedded anywhere)



Datasets

- Citation
- Metadata
- Versioning
- Private URL/*Anonymous Peer Review
- Custom Terms/*Multiple Licenses/Permissions
- Guestbooks
- Publishing Workflows



Files

- Citation
- Ingest
- Preview/Explore
- Metadata/Provenance
- Versioning
- Permissions/Embargo/Restrictions

Subject: Other
License/Data Use Agreement: Custom Dataset Terms

Files Metadata Terms Versions

Search this dataset...

Filter by: File Type: All Access: All File Tag: All

1 to 10 of 18 Files

- GSP_DataUse_Terms_140422.pdf
Author: GSP, 1887-1926
Published: Aug 24, 2014
1,200 Downloads
- GSP_DataUse_Terms_140422.pdf
Author: GSP, 1887-1926
Published: Aug 24, 2014
1,200 Downloads
- GSP_DataUse_Terms_140422.pdf
Author: GSP, 1887-1926
Published: Aug 24, 2014
1,200 Downloads
- GSP_DataUse_Terms_140422.pdf
Author: GSP, 1887-1926
Published: Aug 24, 2014
1,200 Downloads
- GSP_DataUse_Terms_140422.pdf
Author: GSP, 1887-1926
Published: Aug 24, 2014
1,200 Downloads
- GSP_DataUse_Terms_140422.pdf
Author: GSP, 1887-1926
Published: Aug 24, 2014
1,200 Downloads
- GSP_DataUse_Terms_140422.pdf
Author: GSP, 1887-1926
Published: Aug 24, 2014
1,200 Downloads
- GSP_DataUse_Terms_140422.pdf
Author: GSP, 1887-1926
Published: Aug 24, 2014
1,200 Downloads
- GSP_DataUse_Terms_140422.pdf
Author: GSP, 1887-1926
Published: Aug 24, 2014
1,200 Downloads
- GSP_DataUse_Terms_140422.pdf
Author: GSP, 1887-1926
Published: Aug 24, 2014
1,200 Downloads

Edit Files

- Metadata
- Restrict
- Unrestrict
- Tags
- Embargo
- Delete

Democracy Credibility Revisited

id	is3	year	leadcode	portfolio_1	IMFIn	IMFonset	IMFongoing	maleshare_minister	male	female	mingr	genderport	prestport	incid	f
3	AFG	1980	170	Foreign Relations	0	0	0	1	1	0	0	-1	1	2	0
3	AFG	1980	170	Planning & Development	0	0	0	1	1	0	0	0	0	2	0
3	AFG	1980	170	Construction & Public	0	0	0	1	1	0	0	-1	0	2	0

File Tools Metadata Versions

File Tools - Open in New Window

Tell me what you want to know

What is this data about?

Answer please

The data table contains information on various aspects of government and society, such as foreign relations, planning and development, construction and public works, finance, budget and treasury, government, interior and home affairs, health and social welfare, environment, environment and information, transport, and agriculture, food, fisheries and livestock.

id	is3	year	leadcode	portfolio_1	IMFIn	IMFonset	IMFongoing	maleshare_minister	male	female	mingr	genderport	prestport	incid	f
3	AFG	1980	170	Foreign Relations	0	0	0	1	1	0	0	-1	1	2	0
3	AFG	1980	170	Planning & Development	0	0	0	1	1	0	0	0	0	2	0
3	AFG	1980	170	Construction & Public	0	0	0	1	1	0	0	-1	0	2	0

Asset Citation

Spacie, Aimee; Song, Lucy, 2020, "Democracy Credibility Revisited", https://doi.org/10.7910/D1/1_UNF-EM-FKL7sgXtAKW4KwISw== [54kUNF]

File Dataset - Learn about Data Citation Standards.

Preview Metadata Versions

Open in New Window

id	is3	year	leadcode	portfolio_1	IMFIn	IMFonset	IMFongoing	maleshare_minister	male	female	mingr	genderport	prestport	incid	f
1	2	USA	1816	D	0	0	0	3.2							
2	2	USA	1817	D	0	0	0	3.2							
3	2	USA	1818	D	0	0	0	3.2							
4	2	USA	1819	D	0	0	0	3.2							
5	2	USA	1820	D	0	0	0	3.2							
6	2	USA	1821	D	0	0	0	3.2							
7	2	USA	1822	D	0	0	0	3.2							
8	2	USA	1823	D	0	0	0	3.2							
9	2	USA	1824	D	0	0	0	3.2							
10	2	USA	1825	D	0	0	0	3.2							
11	2	USA	1826	D	0	0	0	3.2							
12	2	USA	1827	D	0	0	0	3.2							
13	2	USA	1828	D	0	0	0	3.2							
14	2	USA	1829	D	0	0	0	3.2							
15	2	USA	1830	D	0	0	0	3.2							
16	2	USA	1831	1	1	0	0	3.2							
17	2	USA	1832	1	1	0	0	3.2							
18	2	USA	1833	1	1	0	0	3.2							
19	2	USA	1834	1	1	0	0	3.2							
20	2	USA	1835	1	1	0	0	3.2							
21	2	USA	1836	1	1	0	0	3.2							
22	2	USA	1837	1	1	0	0	3.2							
23	2	USA	1838	1	1	0	0	3.2							
24	2	USA	1839	1	1	0	0	3.2							
25	2	USA	1840	1	1	0	0	3.2							
26	2	USA	1841	1	1	0	0	3.2							
27	2	USA	1842	1	1	0	0	3.2							

The User Experience: Collections

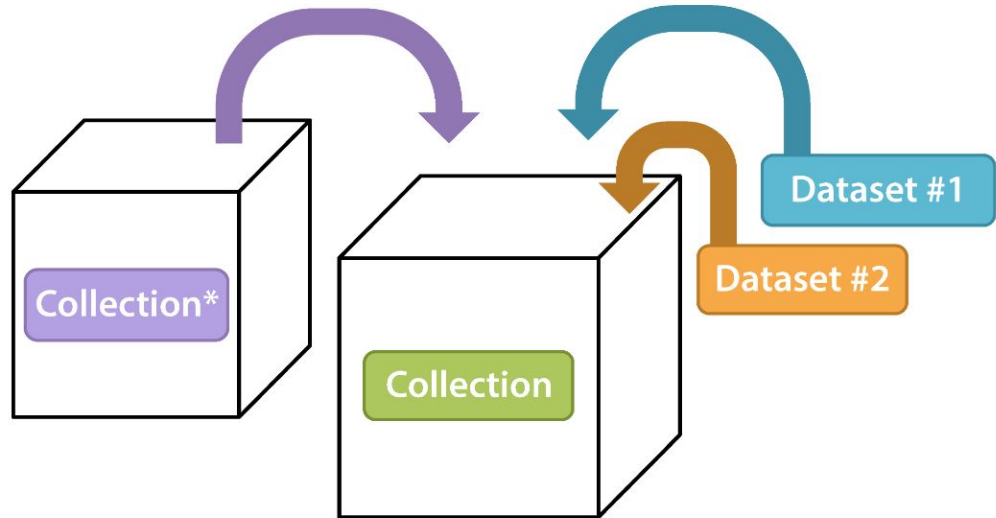
Create a Dataverse

“Collection”

- Edit Dataverse Collection
 - General Information
 - **Theme**
 - Widgets
 - Dataverse Collection Search Box Widget
 - Dataverse Collection Listing Widget
 - Adding Widgets to an OpenScholar Website
 - **Roles & Permissions**
 - Setting Access Configurations
 - Assigning Roles to Users and Groups
 - Dataset **Templates**
 - Dataset **Guestbooks**
 - Featured Dataverse Collection
- Dataset Linking
- Dataverse Collection Linking
- Publish Your Dataverse Collection

A Dataverse collection is a container for datasets (research data, code, documentation, and metadata) and other Dataverse collections, which can be setup for individual researchers, departments, journals and organizations.

Schematic Diagram of a **Collection** in Dataverse Software 5.0



Container for your **Datasets** and/or **Collections***

* Collections can contain other Collections

*contact support@dataverse.harvard.edu and request the creation of a collection

HARVARD
Dataverse

Add Data Search About User Guide Support Sign Up Log In

GARY KING

Gary King Dataverse <http://gking.harvard.edu/>
(Harvard University)

Harvard Dataverse >

Contact Share

Gary King is the Albert J. Weatherhead III University Professor at Harvard University -- one of 25 with Harvard's most distinguished faculty title -- and Director of the Institute for Quantitative Social Science. King develops and applies empirical methods in many areas of social science, focusing on innovations that span the range from statistical theory to practical application.

You can see the related collections for Dr. King here:
 Social Science One
 Gov2001
 Readme2 Replication
 ROAD

Search this dataverse... Advanced Search


Datasets (80)
 Files (2,032)

1 to 10 of 80 Results Sort

Archive of the Controversy Involving Wendy K. Tam Cho, Brian J. Gaines, and the American Political Science Review
 Jan 11, 2024

Burden, Barry C.; Kimball, David C.; King, Gary, 2004, "Archive of the Controversy Involving Wendy K. Tam Cho, Brian J. Gaines, and the American Political Science Review", <https://doi.org/10.7910/DVN/NJ8OZQ>, Harvard Dataverse, V1

GARY KING BIO & C.V. WRITINGS RESEARCH AREAS SOFTWARE DATAVERSE RESEARCH GROUP TEACHING CONTACT



Gary King is the Weatherhead University Professor at Harvard University. He also serves as Director of the Institute for Quantitative Social Science. He and his research group develop and apply empirical methods in many areas of social science research. [Full bio and CV](#)

Research Areas

Methods Applications

- Anchoring Vignettes (for interpersonal incomparability)
- Automated Text Analysis
- Causal Inference
- Event Countdown Divergence
- Evaluating Social Security Forecasts
- Incumbency Advantage
- Chinese Censorship
- Measuring Health Care Production

GARY KING BIO & C.V. WRITINGS RESEARCH AREAS SOFTWARE DATAVERSE RESEARCH GROUP TEACHING CONTACT

HOME /

My Dataverse collection

Below is my Dataverse collection, which is comprised of data sets and replication data sets associated with my published articles and books. For information about the Dataverse project (which I created and run), see [this article](#) and the [Dataverse.org](#) project website.

Harvard Dataverse >

Search this dataverse... Advanced Search

Datasets (80)
 Files (2,032)

1 to 10 of 80 Results Sort

Archive of the Controversy Involving Wendy K. Tam Cho, Brian J. Gaines, and the American Political Science Review
 Jan 11, 2024

Burden, Barry C.; Kimball, David C.; King, Gary, 2004, "Archive of the Controversy Involving Wendy K. Tam Cho, Brian J. Gaines, and the American Political Science Review", <https://doi.org/10.7910/DVN/NJ8OZQ>, Harvard Dataverse, V1

An article by Barry C. Burden and David C. Kimball entitled "A New Approach to the Study of Ticket Spitting" was published in the September 1998 issue of the American Political Science Review1. The empirical part of the article made use of an experiment designed to test the effect of

More...



DANIEL P. ALDRICH

Daniel P. Aldrich is full professor of political science and Director of the [Security and Resilience Studies Program](#) at [Northeastern University](#). He researches post-disaster recovery, countering violent extremism, the siting of controversial facilities and the interaction between civil society and the state. He has held posts as a Fulbright Research Fellow and an Abe Fellow at Tokyo University and as an AAAS Science and Technology Fellow with USAID. Aldrich is a contributor to the New York Times, CNN, The Conversation, and the Asahi Shinbun, among other media.



Daniel Aldrich Dataverse

Northeastern University

(Northeastern University)

Harvard Dataverse >

Contact Share

Search this dataverse...



Advanced Search

 **Dataverses (0)**

 **Datasets (10)**

 **Files (27)**

Publication Year

2022 (1)

2020 (3)

2019 (2)

2016 (1)

2012 (1)

More...

1 to 10 of 10 Results

Sort

Replication Data for: How Social Infrastructure Saves Lives: A Quantitative Analysis of Japan's 3/11 Disasters

Nov 23, 2022



Aldrich, Daniel, 2022, "Replication Data for: How Social Infrastructure Saves Lives: A Quantitative Analysis of Japan's 3/11 Disasters", <https://doi.org/10.7910/DVN/YA0EYS>, Harvard Dataverse, V1, UNF:6:fe9B3jEgPRcE25dCykAt6g== [fileUNF]

Observers have long debated how societies should invest resources to safeguard citizens and property, especially in the face of increasing shocks and crises. This article explores how social infrastructure -the spaces and places that help build and maintain social ties and trust...

Brewing Social Capital: A Survey of Breweries and Social Ties across the United States

Dec 20, 2020




Aldrich, Daniel; Schulz, Eric, 2020, "Brewing Social Capital: A Survey of Breweries and Social Ties across the United States",

陽明交大 NYCU **Dataverse** Add Data Search User Guide

NYCU Dataverse

Metrics 1,741 Downloads

 Wei-Chen (Walton) Chiu Dataverse


Search this dataverse... Advanced Search

Dataverses (1)
Datasets (177)
 Files (4,802)

Dataverse Category
 Researcher (1)


Publication Year

1 to 10 of 178 Results

 **DenseNet-Powered Radiologist Early Warning System by RGB Imaging**
 Jan 31, 2024
 Shyh-Wei Chen; Yu-Heng Hsieh; Jyun-Kai Chen; Shin-Fu Wang; Wei Ching-Tsornng Tsai; Jyh-Wen Chai; Shyan-Ming Yuan, 2024, "DenseNet"
<https://doi.org/10.57770/102YL4C>, NYCU Dataverse, V1

- Author Name**
- Wei-Chen Chiu (28)
 - Tan, Chh Shari (21)
 - Sheng-Hsiung Yang (19)
 - Yau-Huei Wei (15)
 - Y. H. Wei (14)
 - Yu-Hsuan Tsai (13)
 - Min Sun (7)
 - Shyan-Ming Yuan (7)
 - Yu-Ting Wu (7)
 - Yuan, Shyan-Ming (6)
 - Fu-En Wang (4)
 - Jin-Wu Tsai (4)
 - Mario Fritz (4)
 - Y. T. Wu (4)
 - Yu-Ju Tseng (4)
 - Less...
- [More...](#)

陽明交大 NYCU **Dataverse** Add Data Search User Guide About Support English Log In

 Wei-Chen (Walton) Chiu Dataverse Visit my personal website
 (Computer Science, NYCU)

NYCU Dataverse >

Associate Professor in Computer Science at National Chiao Tung University

Frank S Chen Junior Chair Professor cacaFly Junior Chair Professor

Doctor of Engineering Science (Dr.-Ing.) Max Planck Institute for Informatics

Search this dataverse... Advanced Search

Dataverses (0)
Datasets (3)
 Files (1,074)

Publication Year
 2022 (3)

Subject
 Computer and Information Science (2)

Author Affiliation
 Max Planck Institute for Informatics (2)
 Department of Computer Science, National Chiao Tung University (1)

1 to 3 of 3 Results IT Sort

Replication Data for: Using generification to create and label photos that are challenging for computer vision and people
 Jun 14, 2022
 Piotr Kulewki; Xu-Jing Chang; Yeng-Chi-Tsun; Wei-Chen Chiu; Yung-Ju Chang, 2022, "Replication Data for Using generification to create and label photos that are challenging for computer vision and people", <https://doi.org/10.57770/102YL4C>, NYCU Dataverse, V1
 It would be hard to overstate the importance of Computer Vision (CV), applications of which can be found from self-driving cars, through facial recognition to augmented reality and the healthcare industry. Recent years have witnessed dramatic progress in visual-object recognition...

Replication Data for: Improving the Kinect by Cross-Modal Stereo
 Jun 14, 2022
 Wei-Chen Chiu; Uri Barnea; Mario Fritz, 2022, "Replication Data for Improving the Kinect by Cross-Modal Stereo", <https://doi.org/10.57770/102YL4C>, NYCU Dataverse, V1
 Kinect's active sensing strategy is very well suited to produce robust and high-frame rate depth maps for human pose estimation. But the shift to the robotics domain surfaced applications under a wider set of operation conditions than it was originally designed for. We see the same...

Replication Data for: Multi-Class Video Co-Segmentation with a Generative Multi-Video Model
 Jun 14, 2022
 Wei-Chen Chiu; Mario Fritz, 2022, "Replication Data for Multi-Class Video Co-Segmentation with a Generative Multi-Video Model", <https://doi.org/10.57770/102YL4C>, NYCU Dataverse, V1
 Video data provides a rich source of information that is available to us today in large quantities e.g. from online resources. Tasks like segmentation benefit greatly from the analysis of spatio-temporal motion patterns in videos and recent advances in video segmentation has shown...

Wei-Chen (Walton) Chiu; Taiwanese: Khu Uî-Sîn

Home About My Research Contact Us

 Associate Professor in Computer Science at National Chiao Tung University

Frank S Chen Junior Chair Professor
 cacaFly Junior Chair Professor

Doctor of Engineering Science (Dr.-Ing.)
 Max Planck Institute for Informatics

walton [at] cs [dot] nctu [dot] edu [dot] tw
 walton [dot] chiu [at] gmail [dot] com

 Curriculum Vitae  Google Scholar

[Office] Room 526, Engineering Building C, Guangfu Campus, NCTU
 [Telephone] +886-3-5712121 ext. 54730

Information about my research group:
 [Name] Enriched Vision Applications Lab
 [Location] Room 548, Engineering Building C, Guangfu Campus, NCTU



Metadata you can use to describe collections

- Names
- Affiliations
- Identifiers of collection URLs
- Categories
- Emails
- Descriptions

New Dataverse

[Demo Dataverse >](#)

*Asterisks indicate required fields

Host Dataverse ⓘ

Dataverse Name * ⓘ
Identifier * ⓘ
Category * ⓘ
Email * ⓘ
Affiliation ⓘ
Description ⓘ

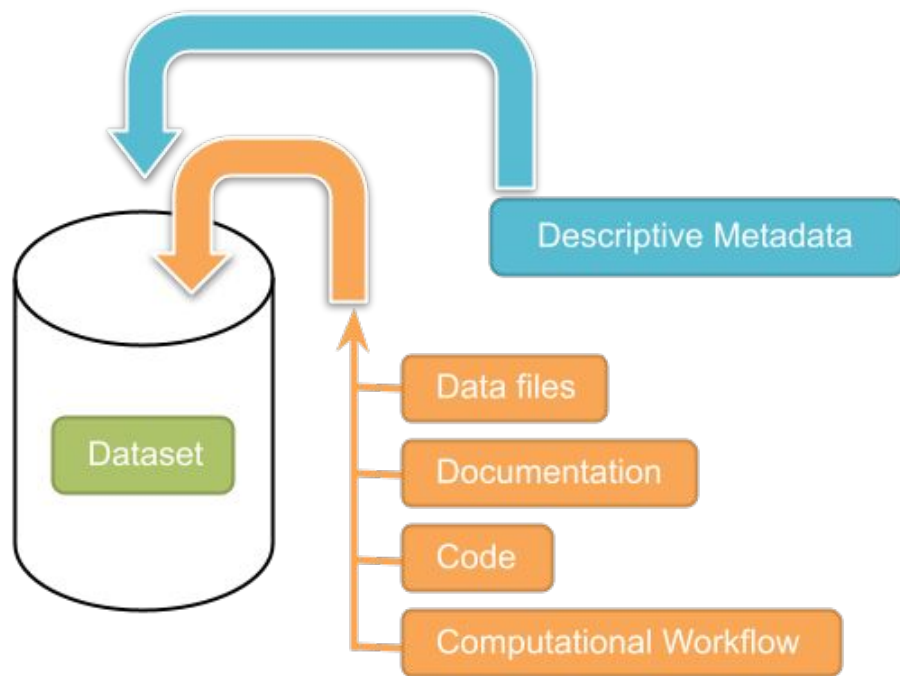
This field supports only certain [HTML tags](#).

The User Experience: Datasets

Deposit a Dataset

- Supported Metadata
- Adding a New Dataset
- File Upload
- File Handling
- Restricted Files
- Edit Files
- Replace Files
- Terms
- Roles & Permissions
- Data Provenance
- Thumbnails + Widgets
- Publish Dataset
- Submit for Review
- Private URL to Review Unpublished Dataset
- Embargoes
- Dataset Versions
 - Version Details
- Dataset Metrics and Make Data Count
- Cloud Storage + Computing
 - Cloud Computing
 - Cloud Storage Access
- Dataset Deaccession

Schematic Diagram of a **Dataset** in Dataverse 4.0



Container for your data, documentation, code, and computational workflow.

Cambodia 2020 Quick Start

(Purdue University Main Campus)

Harvard Dataserve > International Programs in Agriculture (PIA) - Purdue University > Feed the Future Innovation Lab for Food Safety > Cambodia > Cambodia 2020 Quick Start >

Distribution Center Generic E. Coli, Coliform, and Enterobacteriaceae Detection and Quantification Data

Version 1.0



Vipham, Jessica, 2022, "Distribution Center Generic E. Coli, Coliform, and Enterobacteriaceae Detection and Quantification Data", <https://doi.org/10.7910/DVN/0XJFAS>, Harvard Dataserve, V1, UNF:6xY1pY4yKiRyzZkR19Cn3A== [fileUNF]

Cite Dataset - Learn about [Data Citation Standards](#).

Access Dataset -

Contact Owner Share

Dataset Metrics

12 Downloads

Description

The dataset includes data that was collected from tomatoes, cucumbers, and lettuce from a distribution center located in Cambodia. This dataset includes detection and quantification data for all vegetables types, in two seasons (rainy and dry) for generic E. coli, Coliforms, and Enterobacteriaceae.

Subject

Agricultural Sciences

Keyword

Cambodia, Vegetables, Food Safety, Cambodia, vegetables, food safety, tomatoes, cucumbers, lettuce, coliform bacteria, Enterobacteriaceae

License/Data Use Agreement



CC0 1.0

Files Metadata Terms Versions

Search this dataset...

Filter by

File Type: All - Access: All -

Sort -

1 to 2 of 2 Files

Download -

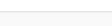
Cambodia_Quantification_codebook.tab

Tabular Data - 581 B

Published Sep 7, 2022

6 Downloads

3 Variables, 7 Observations UNF:6:ati0...CX== codebook for Cambodia quantification data



Cambodia_Quantification_final.tab

Tabular Data - 14.5 KB

Published Sep 7, 2022

6 Downloads

7 Variables, 384 Observations UNF:6:0iN...qGQ==



Harvard Dataserve > Brain Genomics Superstruct Project (GSP) Dataserve >

Brain Genomics Superstruct Project (GSP)

Version 0.0.0

Buckner, Randy L., Roffman, Joshua L., Smoller, Jordan W., 2014, "Brain Genomics Superstruct Project (GSP)", <https://doi.org/10.7910/DVN/025833>, Harvard Dataserve, V10

Cite Dataset - Learn about [Data Citation Standards](#).

Access Dataset -

Contact Owner Share

Dataset Metrics

14,434 Downloads

Description

Large scale imaging data sets are necessary to address complex questions regarding the relationship between brain and behavior. The Brain Genomics Superstruct Project Open Access Data Release exposes a carefully vetted collection of neuroimaging, behavior, cognitive, and personality data for over 1,500 human participants. Each neuroimaging data set includes one high-resolution Magnetic Resonance Imaging (MRI) acquisition and one or more resting-state functional MRI acquisitions. Each functional acquisition is accompanied by a fully-automated quality assessment and pre-computed brain morphometrics are also provided.

Please provide an academic or research institution email address when applying for access.

Subject

Other

License/Data Use Agreement

Custom Dataset Terms

Files Metadata Terms Versions

Search this dataset...

Filter by

File Type: All - Access: All - File Tag: All -

Sort -

1 to 10 of 15 Files

Download

Request Access

GSP_DataUse_Terms_140422.pdf

Adobe PDF - 282.7 KB

Published Aug 24, 2014

1,435 Downloads

MD5: e84...209

Data use terms



GSP_001_140630.csv

Plain Text - 636.4 KB

Published Aug 24, 2014

640 Downloads

MD5: 27e...4e8

Demographic, cognitive/behavior, quality control, and morphometrics data.



GSP_part10_140630.tar

RAW Archive - 10.0 GB

Published May 21, 2014

1,182 Downloads

MD5: 302...48e

Tar archive of imaging data for subjects 1414-1070; refer to GSP_README_140630.pdf for more information.



GSP_part11_140630.tar

RAW Archive - 10.1 GB

Published May 21, 2014

1,229 Downloads

MD5: e02...50e

Tar archive of imaging data for subjects 1-157; refer to GSP_README_140630.pdf for more information.



GSP_part12_140630.tar

RAW Archive - 10.2 GB

Published May 21, 2014

1,254 Downloads

MD5: 0e2...32f

Tar archive of imaging data for subjects 158-314; refer to GSP_README_140630.pdf for more information.



GSP_part13_140630.tar

RAW Archive - 10.1 GB

Published May 21, 2014

1,193 Downloads

MD5: 31e...4ee

Tar archive of imaging data for subjects 315-471; refer to GSP_README_140630.pdf for more information.

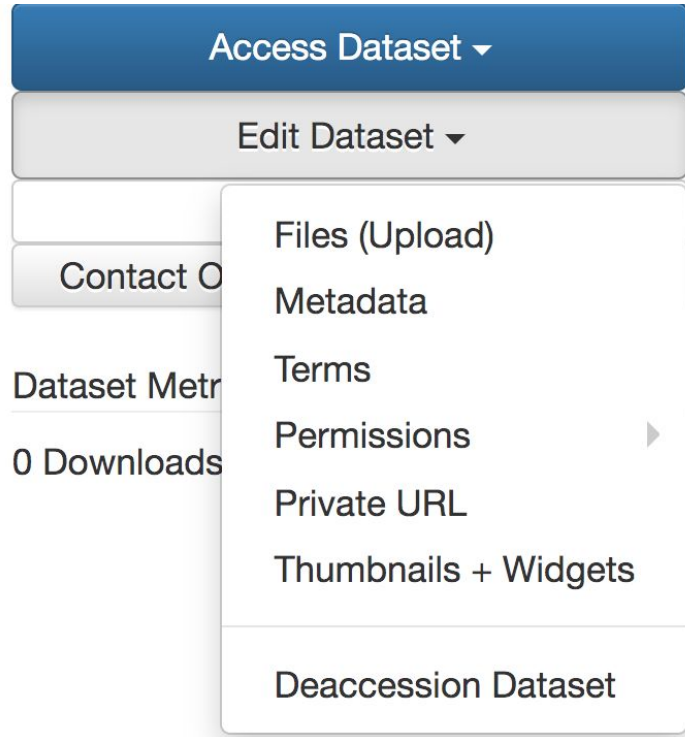


GSP_part14_140630.tar

RAW Archive - 9.7 GB



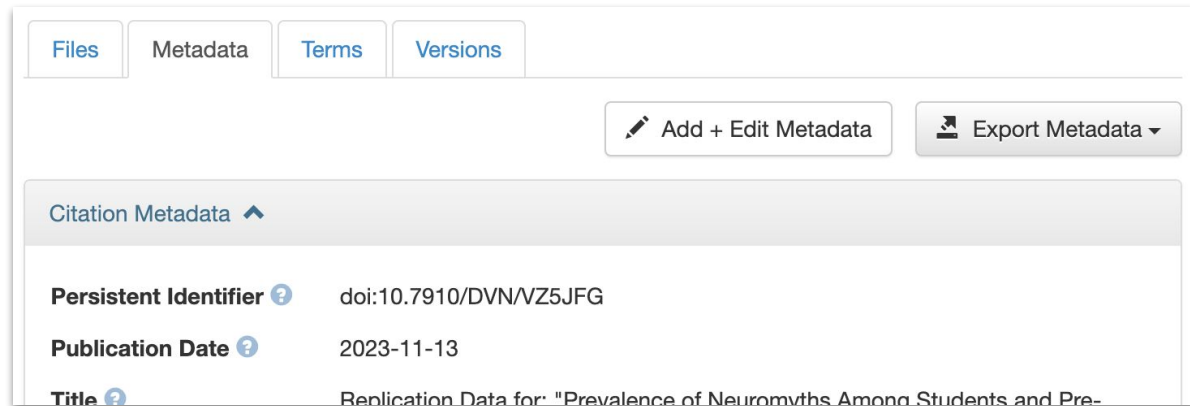
Datasets - Editing Options



- A **Private URL** can be generated for DRAFT datasets to share a dataset with a coauthor or journal, before publishing.

Metadata Dataverse creates for datasets

- Persistent IDs
- Creation dates
- Publication dates
- Version numbers
- Sum of file downloads
- Sum of page views
(in repositories with Make Data Count enabled)



The screenshot displays the metadata interface for a dataset. At the top, there are four tabs: 'Files', 'Metadata', 'Terms', and 'Versions'. Below the tabs are two buttons: 'Add + Edit Metadata' and 'Export Metadata'. The main content area is titled 'Citation Metadata' and contains the following information:

Persistent Identifier ?	doi:10.7910/DVN/VZ5JFG
Publication Date ?	2023-11-13
Title ?	Replication Data for: "Prevalence of Neuromyths Among Students and Pre-

Metadata Dataverse creates for datasets

- Differences among versions

	Dataset Version	Summary	Contributors	
<input type="checkbox"/>	12.0	Citation Metadata: Description (1 Changed); Notes (Changed); Additional Citation Metadata: (3 Changed); Files (Added: 98; Removed: 84); View Details	Julian Gautier	202
<input type="checkbox"/>	11.0	Citation Metadata: Description (1 Changed); Notes (Changed); Additional Citation Metadata: (4 Changed); Files (Added: 84; Removed: 57); View Details	Julian Gautier	202
<input type="checkbox"/>	10.0	Citation Metadata: Notes (Changed); Description (1 Changed); Additional Citation Metadata: (2 Removed, 2 Changed); Files (Added: 57; Removed: 50); View Details	Julian Gautier	202
<input type="checkbox"/>	9.0	Citation Metadata: Description (1 Changed); Additional Citation Metadata: (1 Added); Files (Added: 50; Removed: 37); View Details	Julian Gautier	202
<input type="checkbox"/>	8.1	Citation Metadata: Description (1 Changed); View Details	Julian Gautier	202
<input type="checkbox"/>	8.0	Citation Metadata: Description (1 Changed); Additional Citation Metadata: (1 Added, 1 Changed); Files (Added: 1; Replaced: 1); View Details	Julian Gautier	202
<input type="checkbox"/>	7.0	Citation Metadata: Description (1 Changed); Files (Added: 35; Removed: 36); View Details	Julian Gautier	202

The User Experience: Files

File sizes and upload options and examples

File formats:

- All formats accepted

Per file size limitations*:

- Set by repository

Tabular Ingest*:

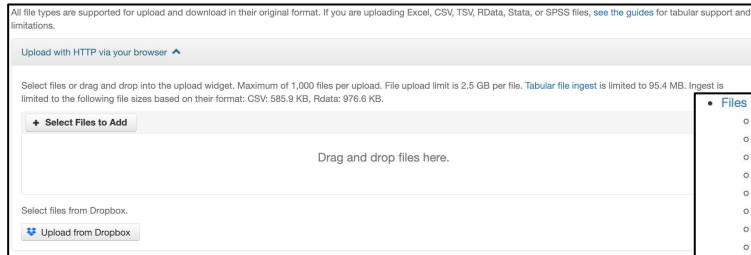
- Set by repository

Total project size limitations*:

- 1 TB (non Harvard)
- 2.5TB (Harvard)
- Paid options available for larger projects

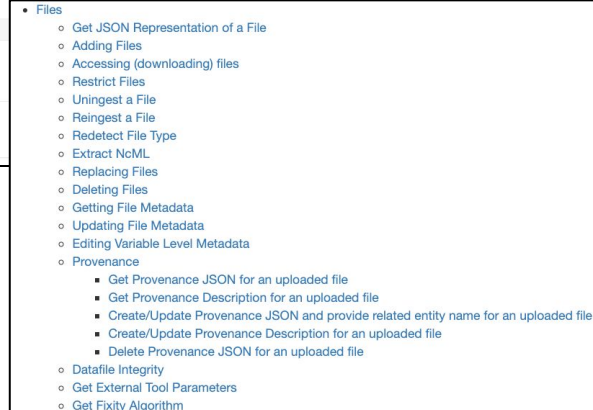
Policies*:

- No identifiable data
- Copyright adherence
- No protected password files
- Terms required (CC0 by default)



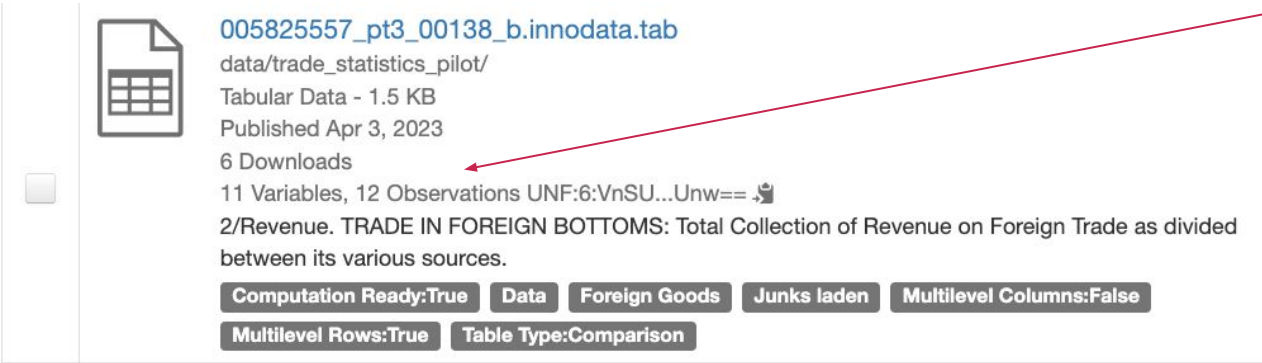
File upload options

Native APIs file handling options



***Determined by repository.** Unless otherwise stated, these policies are for Harvard repository.

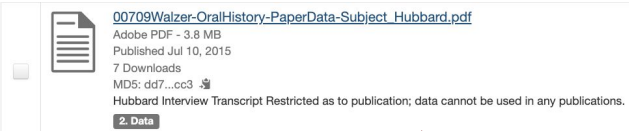
File level metadata considerations and features



005825557_pt3_00138_b.innodata.tab
data/trade_statistics_pilot/
Tabular Data - 1.5 KB
Published Apr 3, 2023
6 Downloads
11 Variables, 12 Observations UNF:6:VnSU...Unw== ↕
2/Revenue. TRADE IN FOREIGN BOTTOMS: Total Collection of Revenue on Foreign Trade as divided between its various sources.

Computation Ready:True Data Foreign Goods Junks laden Multilevel Columns:False
Multilevel Rows:True Table Type:Comparison

Metadata extracted from tabular and other files with tabular ingest feature: **Observations, Variables**



00709Walzer-OralHistory-PaperData-Subject_Hubbard.pdf
Adobe PDF - 3.8 MB
Published Jul 10, 2015
7 Downloads
MD5: dd7...cc3 ↕
Hubbard Interview Transcript Restricted as to publication; data cannot be used in any publications.
2. Data

File level “terms”



005825557_pt3_00140.innodata.txt
data/trade_statistics_pilot/
Plain Text - 194 B
Published Apr 3, 2023
0 Downloads
MD5: e18...5a8 ↕
File associated with data tables series: Population:15

Customs Data JAMES H. HART STATISTICS

File level “tags”

Plan to populate additional metadata for files: description, edit file names, add tags

Towards AI integration

HARVARD
Dataverse

Add Data - Search - About - User Guide - Support - Sign Up - Log In

Harvard Dataverse >

Why insurgents kill civilians in capital cities: A disaggregated analysis of mechanisms and trends

Version 1.0

Koren, Ore, 2017. "Why insurgents kill civilians in capital cities: A disaggregated analysis of mechanisms and trends", <https://doi.org/10.7910/DVN/PHHZI7>, Harvard Dataverse, V1, UNF:6:NDQ2oRK6GGGA+OZgUMdA== [fileUNF]

Cite Dataset - Learn about [Data Citation Standards](#).

Access Dataset -
Contact Owner Share

Dataset Metrics
1,797 Downloads

Description

Research into the causes of violence against civilians has increased significantly in recent years, yet the mechanisms governing spatial patterns of victimization remain poorly understood. My investigation explores if and why one specific locality, capital cities, experiences a higher frequency of violence against civilians perpetrated by armed insurgent organizations. I argue that the political value associated with capitals allows these groups to asymmetrically impose higher costs on the regime by targeting civilians in these localities. I lay out and validate three specific mechanisms to explain this pattern: elite coercion, popular intimidation, and international persuasion. In the first scenario insurgents aim to influence domestic elites directly. In the second, they aim to affect domestic civilians' resolve. In the third, they seek to influence international audiences. Using new geolocated global atrocities data for the years 1996-2009, I evaluate this linkage by employing different methodological approaches and accounting for potential reporting biases. Finally, I show that ethnic and secessionist wars are more likely to experience atrocities in the capital compared with other conflicts. The findings illustrate potential benefits from explaining the temporal and spatial variation in violence by insurgents, with a focus on strategic conditions and power asymmetries. (2017-09-21)

Subject
Social Sciences

Keyword
Political violence; Civilian victimization; Conflict; Spatial analysis

License/Data Use Agreement
 CC0 1.0

Files Metadata Terms Versions

Change View Table Tree

Search this dataset...

Filter by
File Type: All - Access: All - Sort -

21 to 30 of 43 Files Download -

	A.8. Forecasting and Cross Validation Exercises.R Replication files 09_20_17/ R Syntax - 4.9 KB Published Sep 21, 2017 39 Downloads MDS: 9f8...7a5		
	A.9. Capital Atrocities and Regime Survival.R Replication files 09_20_17/ R Syntax - 11.1 KB Published Sep 21, 2017 39 Downloads MDS: 7c1...990		
	eth_sec_war_dat.tab Replication files 09_20_17/Data/Conflict Type Analysis/ Tabular Data - 50.5 MB Published Sep 21, 2017 46 Downloads 48 Variables, 214840 Observations UNF:6:URZ8...MBw==		

<https://doi.org/10.7910/DVN/PHHZI7>



This chatbot only sees the tabular data but is clueless about the metadata

Tell me what this data are about

Harvard Dataverse > Why insurgents kill civilians in capital cities: A disaggregated analysis of mechanisms and trends >

afginc.tab

This file is part of "Why insurgents kill civilians in capital cities: A disaggregated analysis of mechanisms and trends".

Version 1.0

File Citation

Koren, Ore, 2017, "afginc.tab", *Why insurgents kill civilians in capital cities: A disaggregated analysis of mechanisms and trends*, <https://doi.org/10.7910/DVN/PHHZI7/GVJQH3>, Harvard Dataverse, V1, UNF:6:nUddO4V8XEZZCnhI+9rxYg== [fileUNF]

[Cite Data File](#) Learn about [Data Citation Standards](#).

Dataset Citation

Koren, Ore, 2017, "Why insurgents kill civilians in capital cities: A disaggregated analysis of mechanisms and trends", <https://doi.org/10.7910/DVN/PHHZI7>, Harvard Dataverse, V1, UNF:6:NDQ2oRK6QGGrA+OZgUMdDA== [fileUNF]

[Cite Dataset](#) Learn about [Data Citation Standards](#).

File Tools Metadata Versions

File Tools Open in New Window

Tell me what you want to know

What is this data about?

Answer please

This data is about latitude, longitude, and event information.

Unnamed: 0	latitude	longitude	event
1	30.9666666667	61.8833333333	2
2	31	64	3
3	31.0077777778	66.4002777778	3
4	31.0744444444	53.3191666667	1

Access File

Contact Owner

Share

File Metrics


50 Downloads

Cool but poor



This chatbot only sees the tabular data but is clueless about the metadata

Tell me the range of latitudes and longitudes with the highest number of events

 HARVARD
Dataverse

Add Data ▾ Search ▾ About User Guide Support Sign Up Log In

Harvard Dataverse > Why insurgents kill civilians in capital cities: A disaggregated analysis of mechanisms and trends >

afginc.tab

This file is part of "Why insurgents kill civilians in capital cities: A disaggregated analysis of mechanisms and trends".

Version 1.0

File Citation

Koren, Ore, 2017, "afginc.tab", *Why insurgents kill civilians in capital cities: A disaggregated analysis of mechanisms and trends*, <https://doi.org/10.7910/DVN/PHHZ17/GVJOH3>, Harvard Dataverse, V1, UNF:6:nUddO4V8XEZZCnhI+9rxYg== [fileUNF]

[Cite Data File ▾](#) [Learn about Data Citation Standards.](#)

Dataset Citation

Koren, Ore, 2017, "Why insurgents kill civilians in capital cities: A disaggregated analysis of mechanisms and trends", <https://doi.org/10.7910/DVN/PHHZ17>, Harvard Dataverse, V1, UNF:6:NDQ2oRK6QGGrA+OZgUMdDA== [fileUNF]

[Cite Dataset ▾](#) [Learn about Data Citation Standards.](#)

File Tools **Metadata** Versions

File Tools ▾ [Open in New Window](#)

Tell me what you want to know

longitudes with the highest number of events

Answer please

The range of latitudes and longitudes with the highest number of events is 30.9666666666667, 61.8833333333333 to 32.0, 67.25.


Unnamed: 0	latitude	longitude	event
1	30.96666666667	61.88333333333	2
2	31	64	3
3	31.00777777778	66.40027777778	3
4	31.07444444444	53.31916666667	1

ok-ish



This chatbot only sees the tabular data but is clueless about the metadata

Map the range of latitudes and longitudes with the highest number of events to the names of countries

 **HARVARD**
Dataverse

Add Data ▾ Search ▾ About User Guide Support Sign Up Log In

Harvard Dataverse > Why insurgents kill civilians in capital cities: A disaggregated analysis of mechanisms and trends >

afginc.tab

This file is part of "Why insurgents kill civilians in capital cities: A disaggregated analysis of mechanisms and trends".

Version 1.0

File Citation

Koren, Ore, 2017, "afginc.tab", *Why insurgents kill civilians in capital cities: A disaggregated analysis of mechanisms and trends*, <https://doi.org/10.7910/DVN/PHHZ17/GVJOH3>, Harvard Dataverse, V1, UNF:6:nUddO4V8XEZZCnhl+9rxYg== [fileUNF]

[Cite Data File ▾](#) [Learn about Data Citation Standards.](#)

Dataset Citation

Koren, Ore, 2017, "Why insurgents kill civilians in capital cities: A disaggregated analysis of mechanisms and trends", <https://doi.org/10.7910/DVN/PHHZ17>, Harvard Dataverse, V1, UNF:6:NDQ2oRK6QGGrA+OZgUMdDA== [fileUNF]

[Cite Dataset ▾](#) [Learn about Data Citation Standards.](#)

File Tools [Metadata](#) [Versions](#)

File Tools ▾ [Open in New Window](#)

Tell me what you want to know

map the range of latitudes and longitudes with the highest number of events to the names of countries

Answer please

The countries with the highest number of events in the given range of latitudes and longitudes are Afghanistan, Iran, and Pakistan.

Unnamed: 0	latitude	longitude	event
1	30.9666666667	61.8833333333	2
2	31	64	3
3	31.0077777778	66.4002777778	3
4	31.0744444444	53.3191666667	1

LLM kicks in



This chatbot only sees the tabular data but is clueless about the metadata

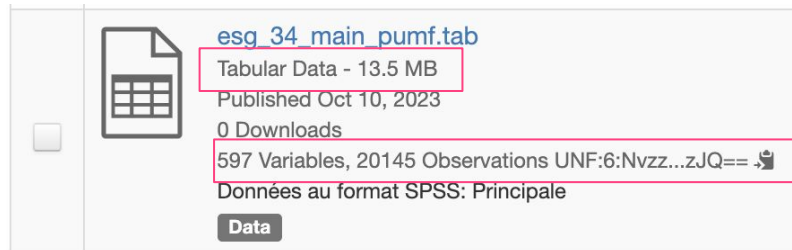
Tabular Data Ingest

Tabular data formats (CSV, Excel, SPSS, Stata, R)

- Tabular data ingest
 - Conversion to Tab-delimited text
 - Metadata extraction of variables
 - Data Documentation Initiative (DDI) metadata standard

Reuses and integrations

- Enables variable metadata searching
- Preview and exploration
- Summary statistics and analysis
- Download as original format, Tab-delimited or R format



The screenshot shows a data file listing interface. On the left is a document icon with a grid. To its right, the file name 'esg_34_main_pumf.tab' is displayed in blue. Below the name, 'Tabular Data - 13.5 MB' is shown in a red box. Further down, 'Published Oct 10, 2023' and '0 Downloads' are listed. A red box highlights the text '597 Variables, 20145 Observations UNF:6:Nvzz...zJQ==', followed by a download icon. Below this, the text 'Données au format SPSS: Principale' is visible. At the bottom of the entry is a dark button labeled 'Data'.

Tabular data: depositing files

- Upload tabular files (SPSS, STATA, R, CSV, Excel)
- Data ingest tips:
 - Responses/ observations as rows
 - Variables/concepts as columns
 - No notes / improper formatting
 - Does not support multiple spreadsheets in one file
 - Excel (XLSX) is supported
 - CSV requires documentation
 - Encode in original language
 - Ingest errors can occur

	v1	v2	v3	v4	v5	v6	v7	v8	v9	v10	v11	v12	v13	v14	v15	v16
1	1.040e+14	4	26	104012	99	1	1973	2	1642	1040	1040	1	6	1	5	5
2	1.040e+14	1	19	104008	99	2	1943	1	99	99	99	1	1	6	4	5
3	1.040e+14	4	98	104003	99	1	1990	1	99	99	99	1	6	3	99	11
4	1.040e+14	2	98	104010	99	1	1983	2	1756	1040	1040	1	6	3	99	11
5	1.040e+14	1	18	104007	99	2	1927	1	99	99	99	1	1	6	4	1
6	1.040e+14	2	19	104007	99	1	1983	1	99	99	99	1	6	2	5	3
7	1.040e+14	4	15	104005	99	2	1970	1	99	99	99	1	1	2	4	6
8	1.040e+14	4	19	104006	99	1	1942	1	99	99	99	2	1	6	5	88
9	1.040e+14	4	15	104005	99	1	1965	2	1040	1276	1040	7	4	1	5	5
10	1.040e+14	7	15	104004	99	2	1955	1	99	99	99	1	1	2	5	5
11	1.040e+14	7	15	104003	99	2	7777	1	99	99	99	1	1	6	77	77
12	1.040e+14	4	18	104005	99	2	1938	1	99	99	99	1	3	6	5	5
13	1.040e+14	7	17	104005	99	1	1945	1	99	99	99	1	1	6	5	4
14	1.040e+14	4	18	104005	99	2	1949	2	1040	1380	1040	1	4	6	5	5
15	1.040e+14	2	15	104003	99	2	7777	1	99	99	99	1	1	2	4	5
16	1.040e+14	4	32	104007	99	1	1974	1	99	99	99	1	6	2	5	5
17	1.040e+14	4	98	104006	99	1	7777	1	99	99	99	1	6	3	99	11
18	1.040e+14	4	25	104010	99	2	1968	1	99	99	99	1	1	2	5	1
19	1.040e+14	4	40	104007	99	1	1967	1	99	99	99	1	1	2	6	2
20	1.040e+14	1	23	104004	99	1	1932	1	99	99	99	1	1	6	4	3
21	1.040e+14	2	18	104003	99	2	1965	1	99	99	99	1	6	2	5	5
22	1.040e+14	1	27	104011	99	1	1956	1	99	99	99	2	1	1	5	2
23	1.040e+14	4	18	104004	99	2	1923	1	99	99	99	1	3	6	3	3
24	1.040e+14	1	19	104006	99	2	1952	1	99	99	99	1	4	2	5	3
25	1.040e+14	3	16	104004	99	2	1947	2	1276	1040	1040	1	1	6	3	3

Tabular data: variable metadata

- Assigns Unique identifiers
- Structured metadata for names, labels, codes, values
- Data contents:
 - Variables
 - Questionnaires & interviews (question text, universes)
 - Recodes, codes, values and notes
- Weighting
- Summary statistics
- Export variable metadata
- Export HTML and XML Codebook

Variable Description
Variable Groups
<ul style="list-style-type: none">• CAN: Cannabis• DV: Derived Variables• ALC: Alcohol• IU: Initial use• VAP: Vaping• OTP: Other tobacco product status• RD: Reference date• DEM2: Demographics 2• SRV: Survey Related Variables• GDR: Gender• TBC: Tobacco
CAN: Cannabis
Variables within CAN: Cannabis
<ul style="list-style-type: none">• Smoked cannabis - lifetime• Age first time smoked cannabis• Frequency smoked cannabis - past 30 days• Nb of days smoked cannabis at least once past month - grouped• Frequency smoked cannabis/tobacco mix - past 30 days• Nb of days smoked can /toba_mix at least once a week - past 30 days• Frequency consumed edibles - past 30 days• Vaped cannabis - lifetime• Age first time vaped cannabis - grouped• Frequency vaped cannabis - past 30 days• Nb of days vaped cannabis at least once past month - 30 days - grouped• Where get devices/liquids - Make your own vaping liquid• Where get devices/liquids - Compassion club/dispensary/storefront• Where get devices/liquids - Online• Where get devices/liquids - Shared around a group of friends• Where get devices/liquids - An acquaintance• Where get devices/liquids - A family member• Where get devices/liquids - A friend• Where get devices/liquids - A dealer• Where get devices/liquids - Other
DV: Derived Variables
Variables within DV: Derived Variables
<ul style="list-style-type: none">• Age group of person• Province of residence (collection)• Smoking Status• Cannabis smoked in the last 30 days• Vaped in the last 30 days• Vaped cannabis in the last 30 days• Consumed alcohol in the last 30 days• First product tried

Source: Canadian Tobacco and Nicotine Survey, 2021.
DDI Codebook (Odesi)

Geospatial Data



[cd_2006_orig_CBF.zip](#)

Shapefile as ZIP Archive - 27.2 MB

Published Apr 4, 2023

0 Downloads

MD5: c0e...22e

UNI-CEN Digital Boundary File for Census Division (CD), 2006, in Esri Shapefile format and NAD83 CSRS / EPSG:3348 projection. This file is part of the CBF-Original Shorelines series. Retrieved from: <https://borealisdata.ca/dataverse/unicen>. For more information about the project, visit <https://observatory.uwo.ca/unicen>.



[pr_1871_orig_CBF.geojson](#)

GeoJSON - 21.5 MB

Published Apr 4, 2023

2 Downloads

MD5: 14f...fa6

UNI-CEN Digital Boundary File for Province/Territory (PR), 1871, in geojson format and WGS84 / EPSG:4326 projection. This file is part of the CBF-Original Shorelines series. Retrieved from: <https://borealisdata.ca/dataverse/unicen>. For more information about the project, visit <https://observatory.uwo.ca/unicen>.



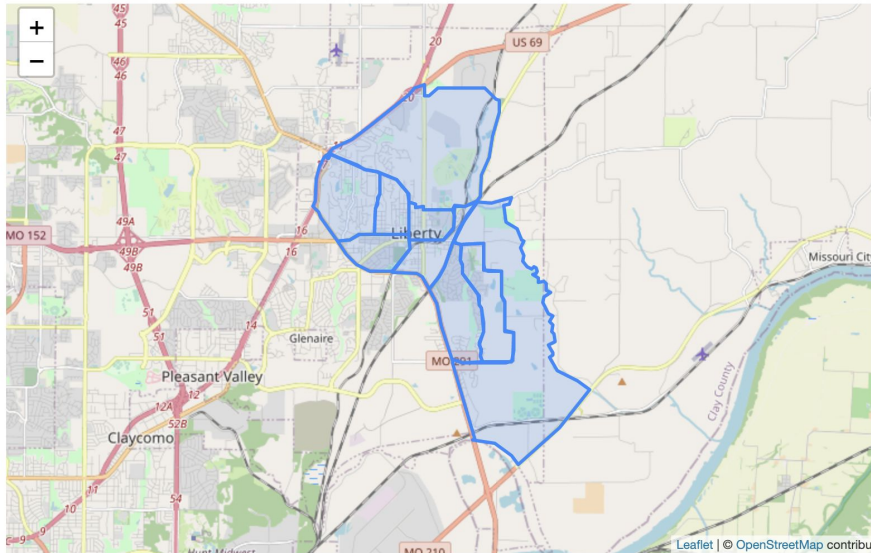
GeoJSON Previewer

Preview

Metadata

Versions

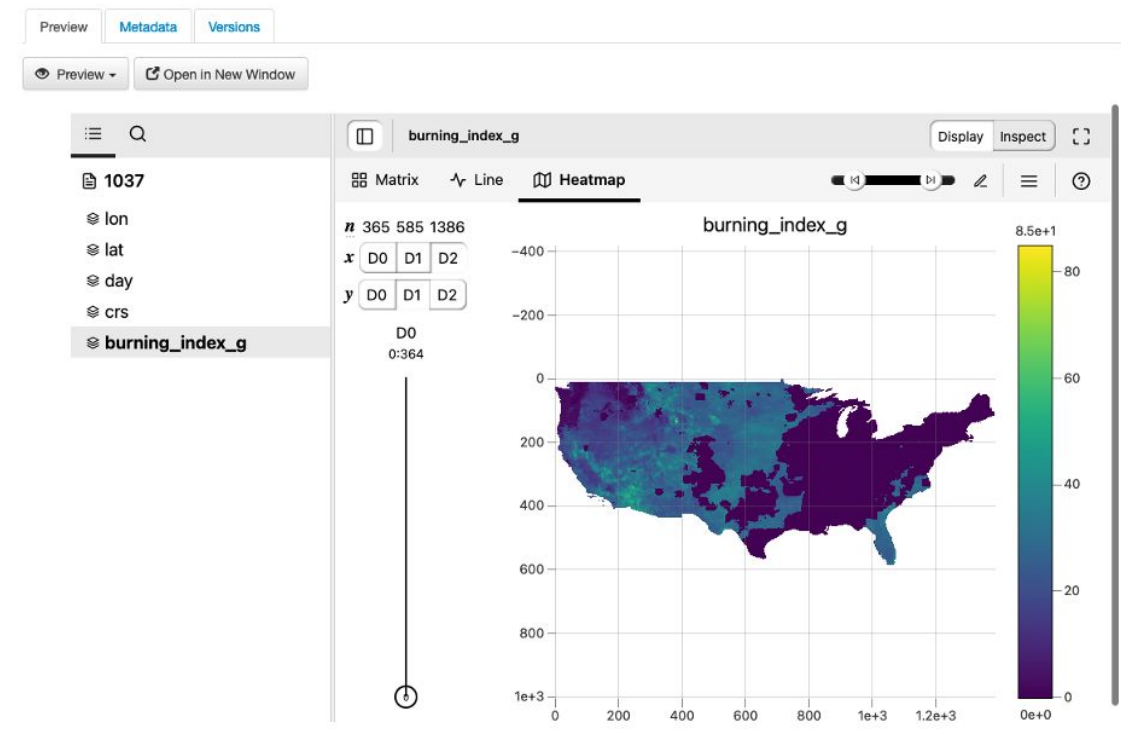
Open in New Window



Previewers originally developed by QDR and maintained at <https://github.com/GlobalDataVerseCommunityConsortium/dataverse-previewers>. Feedback and contributions welcome.

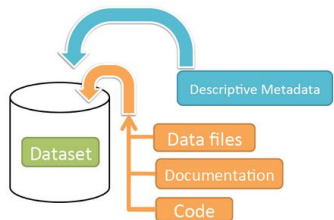
NetCDF & HDF5 Previewer

- Common scientific data formats
- Metadata extracted in NcML (XML) format and saved as an auxiliary file
- H5Web visualization and file preview tool
 - Preview data
 - Explore and visualize data



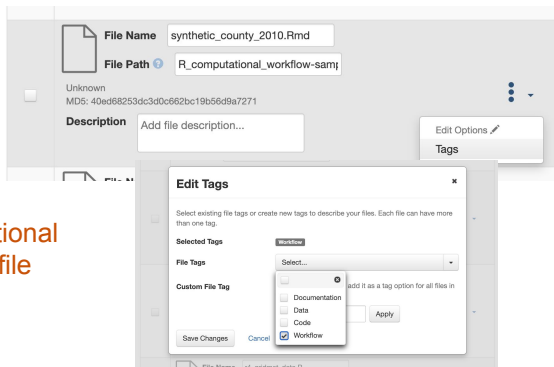
File Types, Format, Documentation

Schematic Diagram of a Dataset in Dataverse 4.0



Container for your data, documentation, and code.

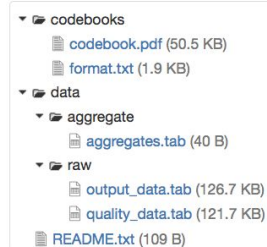
Computational workflow file support



Folder hierarchy support

Change View

Table Tree



Tabular data support

Supported File Formats

Tabular Data ingest supports the following file formats:

File format	Versions supported
SPSS (PDR and SAV formats)	7 to 22
STATA	4 to 15
R	up to 3
Excel	XLSX only (XLS is NOT supported)
CSV (comma-separated values)	(limited support)

File level access control

Access

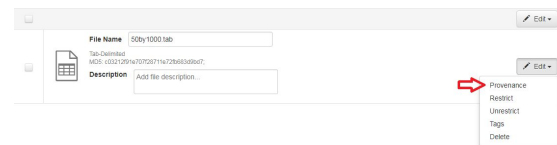
Public (1,784,339)

Restricted (50,591)

Embargoed then Public (154)

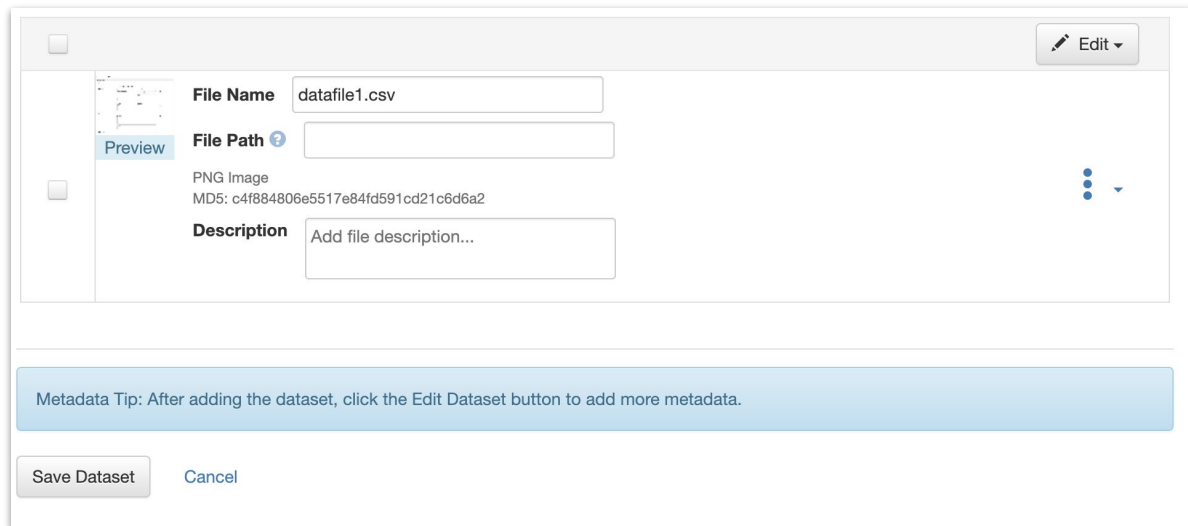
Embargoed then Restricted (14)

File level Provenance support



Metadata you can use to describe files

- File names
- File paths
- Descriptions



The screenshot shows a web interface for adding metadata to a dataset. It features a sidebar with a 'Preview' button and a main form area. The form contains the following fields:

- File Name:** A text input field containing 'datafile1.csv'.
- File Path:** A text input field with a question mark icon.
- File Type:** A label 'PNG Image'.
- MD5:** A text input field containing the hash 'c4f884806e5517e84fd591cd21c6d6a2'.
- Description:** A text input field containing the placeholder text 'Add file description...'.

At the top right of the form is an 'Edit' button with a pencil icon. At the bottom of the form is a blue tip box that reads: 'Metadata Tip: After adding the dataset, click the Edit Dataset button to add more metadata.' Below the tip box are two buttons: 'Save Dataset' and 'Cancel'.

Metadata you can use to describe files

- Tags
- Restrictions
- Embargo dates and reasons

Restrict Access ✕

Restricting limits access to published files. People who want to use the restricted files can request access by default. **If you disable request access, you must add information about access to the Terms of Access field.**

Learn about restricting files and dataset access in the [User Guide](#).

Request Access ? Enable access request

Terms of Access for Restricted Files ?

Metadata you can use to describe files

- Tags
- Restrictions
- Embargo dates and reasons

Edit Embargo ✕

Edit the planned embargo for the selected file or files. Once this dataset version is published, you will need to contact an administrator to change the embargo date or reason of the file or files.

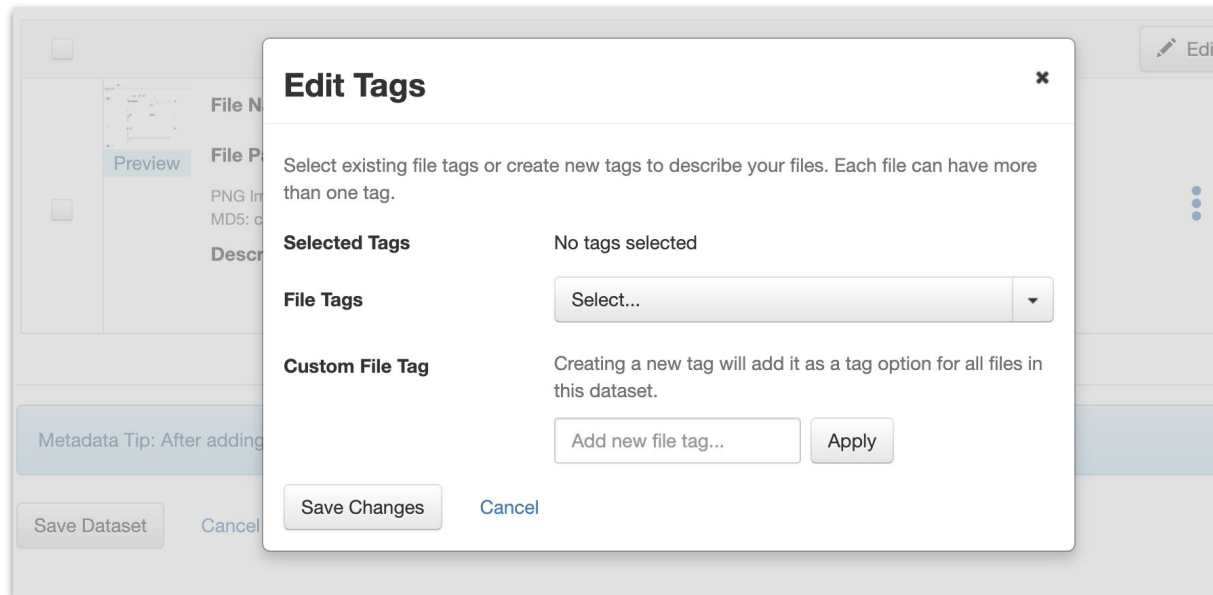
Add or Change **Select the embargo end-date ***

November 2023

S	M	T	W	T	F	S
			1	2	3	4
5	6	7	8	9	10	11
12	13	14	15	16	17	18

Metadata you can use to describe files

- Tags
- Restrictions
- Embargo dates and reasons



Metadata Dataverse creates for ingested tabular files

- Counts of variables and observations
- Variable statistics
- Checksums (UNFs)

File Metadata ^	
File UNF	UNF:6:HSeMI8h1XrpAPAOzE50Qow==
Original File MD5	05ed7a1c7330df3e88e57629bf15b5ff
Type	Tab-Delimited
Variables	8
Observations	1121

TIME

f7570319 Location: Summary Statistics: Min. 2000.0; Valid 1121.0; Mean 2010.6030330062445; Max. 2021.0; StDev 6.318365357642705;

Variable Format: numeric

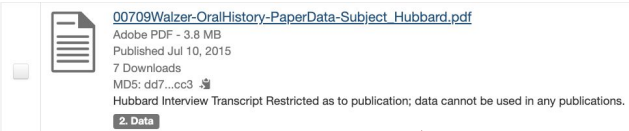
File level metadata considerations and features



005825557_pt3_00138_b.innodata.tab
data/trade_statistics_pilot/
Tabular Data - 1.5 KB
Published Apr 3, 2023
6 Downloads
11 Variables, 12 Observations UNF:6:VnSU...Unw== ↕
2/Revenue. TRADE IN FOREIGN BOTTOMS: Total Collection of Revenue on Foreign Trade as divided between its various sources.

Computation Ready:True Data Foreign Goods Junks laden Multilevel Columns:False
Multilevel Rows:True Table Type:Comparison

Metadata extracted from tabular and other files with tabular ingest feature: **Observations, Variables**



00709Walzer-OralHistory-PaperData-Subject_Hubbard.pdf
Adobe PDF - 3.8 MB
Published Jul 10, 2015
7 Downloads
MD5: dd7...cc3 ↕
Hubbard Interview Transcript Restricted as to publication; data cannot be used in any publications.

2. Data

File level “terms”



005825557_pt3_00140.innodata.txt
data/trade_statistics_pilot/
Plain Text - 194 B
Published Apr 3, 2023
0 Downloads
MD5: e18...5a8 ↕
File associated with data tables series: Population:15

Customs Data JAMES H. HART STATISTICS

File level “tags”

Plan to populate additional metadata for files: description, edit file names, add tags

The User Experience: Files and External Tools

External Tools

- External tools can provide additional features that are not part of the Dataverse Software itself, such as data file previews, visualization, and curation
- Communicate with Dataverse via its robust APIs
- Can be defined for Datasets or Files
- Can be Read-only or Read / Write

Contents:









- [Inventory of External Tools](#)
- [Managing External Tools](#)
 - [Adding External Tools to a Dataverse Installation](#)
 - [Listing All External Tools in a Dataverse Installation](#)
 - [Showing an External Tool in a Dataverse Installation](#)
 - [Removing an External Tool From a Dataverse Installation](#)
- [Testing External Tools](#)
 - [File Level vs. Dataset Level](#)
 - [File Level Explore Tools](#)
 - [File Level Preview Tools](#)
 - [File Level Query Tools](#)
 - [File Level Configure Tools](#)
 - [Dataset Level Explore Tools](#)
 - [Dataset Level Configure Tools](#)
- [Writing Your Own External Tool](#)

File Previewers

- A set of tools that display the content of files, allowing them to be viewed without downloading the file, including
 - audio
 - html
 - Hypothes.is annotations
 - images
 - PDF
 - text
 - video
 - tabular data
 - spreadsheets

1 to 2 of 2 Files

Edit Files ▾ Download

<input type="checkbox"/>		argentina.jpeg JPEG Image - 76.8 KB Deposited Oct 10, 2023 MD5: cec...086	  
<input type="checkbox"/>		messi.jpeg JPEG Image - 12.2 KB Deposited Oct 10, 2023 MD5: 384...6c2	  

Preview Metadata Versions

Explore on View Image



File Previewers (more examples)

Preview Metadata Versions

Explore on Read Document

Previous Next Page: 13 / 59

GEOMETRICAL SOLUTIONS DERIVED FROM MECHANICS.

ARCHIMEDES TO ERATOSTHENES, GREETING:

Some time ago I sent you some theorems I had discovered, writing down only the propositions because I wished you to find their demonstrations which had not been given. The propositions of the theorems which I sent you were the following:

1. If in a perpendicular prism with a parallelogram¹ for base a cylinder is inscribed which has its bases in the opposite parallelograms¹ and its surface touching the other planes of the prism, and if a plane is passed through the center of the circle that is the base of the cylinder and one side of the square lying in the opposite plane, then that plane will cut off from the cylinder a section which is bounded by two planes, the intersecting plane and the one in which the base of the cylinder lies, and also by as much of the surface of the cylinder as lies between these same planes: and the detached

File Tools Metadata Versions

File Tools Open in New Window

	NameFootballPlayer	International
1	Pedro Bonifacio Suárez Pérez \Arico\ (Pedro Suárez)	Argentina
2	Milovan \El Grande Milovan\ Jakšić	Yugoslavia
3	Ernest Libérati	France
4	Alexandre Villaplane (Captain)	France
5	Roberto Gayón Márquez	Mexico
6	Andrew \Andy\ Auld	United States
7	James \Jim\ Brown	United States
8	Jimmy Gallagher	United States
9	Alfred Eisenbeisser	Romania
10	Ladislau Raffinsky	Romania
11	Bartholomew \Bertie\ or \Bart\ McGhee (Bart McGhee)	United States
12	George Moorhouse	United States
13	Alexander "Alec" Wood	United States
14	Lorenzo Fernández	Uruguay
15	Constantino Urbieto Sosa	Argentina
16	Štefan Čambal	Czechoslovakia
17	Ferdinand Daučík	Czechoslovakia
18	Géza Kalocsay	Czechoslovakia
19	František Svoboda	Czechoslovakia
20	Matthias Sindelar (born as Matěj Šindelář)	Austria
21	Joseph Alcazar	France
22	Roger Courtois	France

Zip File Previewer +

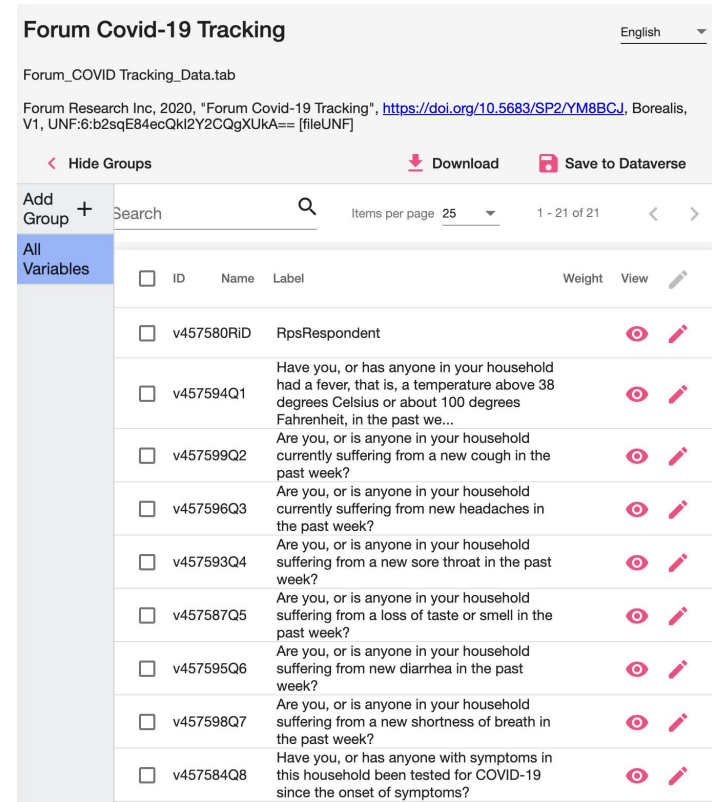
- A previewer that will show you the internal content and structure of a zip file (or electronic lab notebook)
- Uses the Range functionality in our Access api; so it's not just a viewer, it's an individual file unpacker

The screenshot shows a web interface for a zip file previewer. At the top, there are three tabs: "Preview" (selected), "Metadata", and "Versions". Below the tabs is a button labeled "Open in New Window". A message states: "To download the complete zip file, please use the **Access File** button above." Below this is a table showing the contents of a zip file named "BalsacGallery".

File Name	Size	Download Link
BalsacGallery		
BalG_Adsorbates		
Ag110EtO.eps	44.26 kB	Download
Ag110EtO.jpg	147.77 kB	Download
C2H4-C2H22.gif	98.23 kB	Download
COO-CO2.gif	49.64 kB	Download
Cu100KMr.eps	41.86 kB	Download
Cu100KMr.jpg	145.94 kB	Download
Cu100LiMr.eps	43.06 kB	Download
Cu100LiMr.jpg	142.56 kB	Download
Cu100NaMr.eps	41.86 kB	Download
Cu100NaMr.jpg	145.92 kB	Download
Cu110NHxO.eps	24.66 kB	Download
Cu110NHxO.jpg	127.20 kB	Download
Cu111+(4x4)-C60.eps	373.27 kB	Download
Cu111+(4x4)-C60.jpg	379.56 kB	Download
Cu111-C9H10allyl.eps	54.77 kB	Download
Cu111-C9H10allyl.jpg	154.49 kB	Download
Cu111-C9H10alpha.eps	54.78 kB	Download
Cu111-C9H10alpha.jpg	159.73 kB	Download
Cu111-C9H10trans.eps	54.78 kB	Download
Cu111-C9H10trans.jpg	160.98 kB	Download

File Exploration, Curation, and Query Tools

- File level **explore** tools provide a variety of features from data visualization to statistical analysis
- File level **configure** tools allow (authorized) users to send metadata about the file back to Dataverse
- File level **query** tools allow the user to ask questions (e.g. natural language queries) of a data table's contents without having to download the file

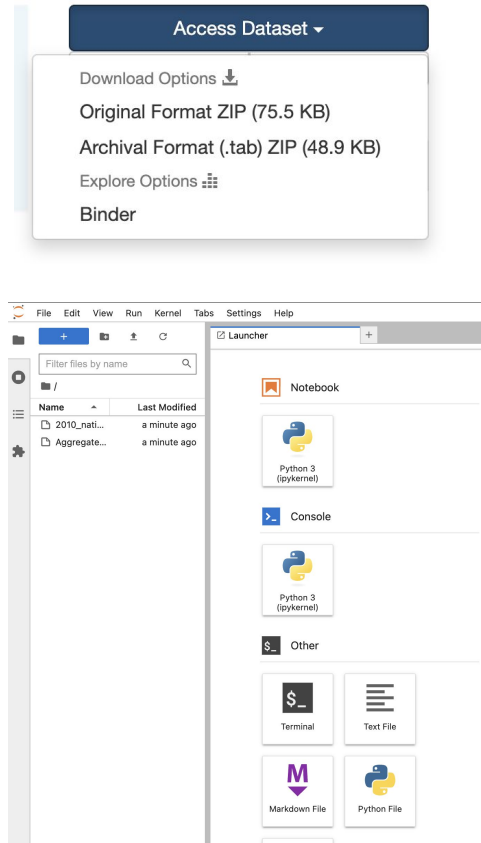


The screenshot displays the 'Forum Covid-19 Tracking' interface. At the top, it shows the dataset name 'Forum_COVID Tracking_Data.tab' and a description: 'Forum Research Inc, 2020, "Forum Covid-19 Tracking", <https://doi.org/10.5683/SP2/YM8BCJ>, Borealis, V1, UNF:6:b2sqE84ecQkl2Y2CQgXUkA== [fileUNF]'. Below the description are buttons for 'Hide Groups', 'Download', and 'Save to Dataverse'. The main area is a table of variables with columns for 'ID', 'Name', 'Label', 'Weight', and 'View'. The 'All Variables' tab is selected on the left. The table lists several variables, each with a checkbox, an ID, a name, a label, and icons for viewing and editing.

<input type="checkbox"/>	ID	Name	Label	Weight	View	
<input type="checkbox"/>	v457580RID	RpsRespondent				
<input type="checkbox"/>	v457594Q1		Have you, or has anyone in your household had a fever, that is, a temperature above 38 degrees Celsius or about 100 degrees Fahrenheit, in the past we...			
<input type="checkbox"/>	v457599Q2		Are you, or is anyone in your household currently suffering from a new cough in the past week?			
<input type="checkbox"/>	v457596Q3		Are you, or is anyone in your household currently suffering from new headaches in the past week?			
<input type="checkbox"/>	v457593Q4		Are you, or is anyone in your household suffering from a new sore throat in the past week?			
<input type="checkbox"/>	v457587Q5		Are you, or is anyone in your household suffering from a loss of taste or smell in the past week?			
<input type="checkbox"/>	v457595Q6		Are you, or is anyone in your household suffering from new diarrhea in the past week?			
<input type="checkbox"/>	v457598Q7		Are you, or is anyone in your household suffering from a new shortness of breath in the past week?			
<input type="checkbox"/>	v457584Q8		Have you, or has anyone with symptoms in this household been tested for COVID-19 since the onset of symptoms?			

Dataset External Tools

- Dataset level **explore** tools allow the user to explore all the files in a dataset - common use case is reproducibility
 - **WholeTale** - creates reproducible research packages based on popular tools such as Jupyter and RStudio
 - **Binder** - spins up custom computing environments in the cloud (including Jupyter notebooks)
- Dataset level **configure** tools allow (authorized) users to send metadata about the dataset back to Dataverse
 - **Turbo Curator** (*coming soon*) - provides recommendations to improve the dataset metadata

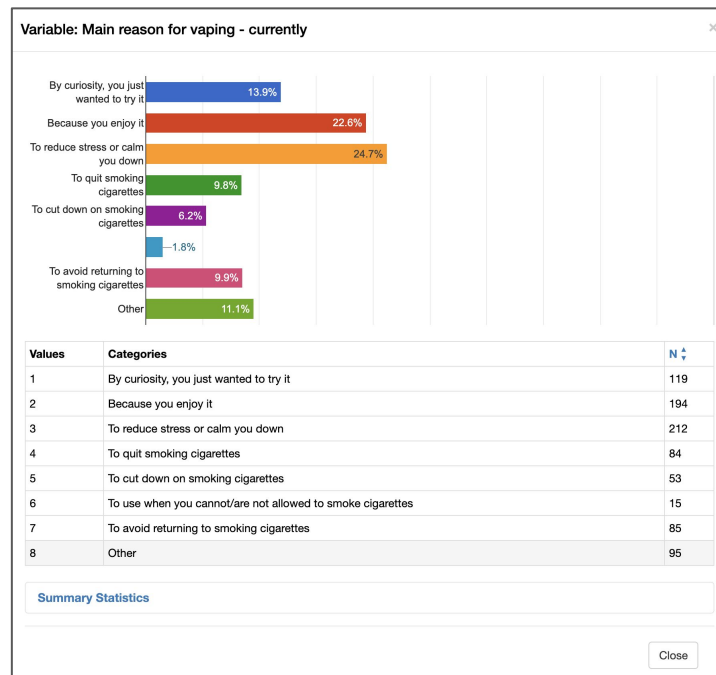


Spotlight: Survey Data Analysis using Data Explorer

- Data Explorer 2.0 [Github](#) (open-source external Dataverse tool)
- Contributed by Borealis (Canada)

Data Analysis using Data Explorer

- Helps users understand the data
 - Preview data and metadata for variables, questions, notes, universes
 - Get started with data analysis; digging, filtering, summary statistics
- Analyse and explore data
 - Frequencies charts
 - Cross-tabulation builder
 - Subsetting & download
- Analysing in Data Explorer promotes:
 - Data reuse
 - Transparency and reproducibility
 - Teaching data concepts
 - Open access and sharing



Source: VAP_35R, Canadian Tobacco and Nicotine Survey, 2021.
Odesi <https://odesi.ca>

Connecting to Data Explorer



The screenshot shows a list of four files in a web interface. A context menu is open over the 'Data Explorer' option, which is highlighted with a red box. The files listed are:

- CTNS-5305-E-2021_F1.tab**: Tabular Data - 2.4 MB, Published Apr 15, 2023, 0 Downloads. Description: 104 Variables, 9908 Observations UNF:6:15C]...dpA== Microdata in tab separated value format.
- CTNS2021_AbridgedUserGuide.pdf**: Survey documentation/ Adobe PDF - 929.6 KB, Published Apr 15, 2023, 0 Downloads. Description: MDS: 717...891 Data Documentation.
- CTNS2021_PUMF_Cdbk.pdf**: Survey documentation/ Adobe PDF - 1.4 MB, Published Apr 15, 2023, 0 Downloads. Description: MDS: fa9...614 Data Documentation.
- CTNS2021_Questionnaire_E.pdf**: Survey documentation/ Adobe PDF - 741.6 KB, Published Apr 15, 2023, 0 Downloads.

The context menu options are:

- File Access
- Public
- Download Options
- SPSS Binary (Original File Format)
- Tab-Delimited
- RData
- Download Metadata
- Variable Metadata
- Data File Citation
- Explore Options
- View Data
- Data Explorer** (highlighted)

Data Explorer

Search for variables

Cross tabulate two or more variables

View multiple summary statistics

Subset and download

Employment Insurance Coverage Survey, 2021 [Canada]

EICS-4428-E-2021_F1.tab

Labour Statistics Division, 2023, "Employment Insurance Coverage Survey, 2021 [Canada]" <https://doi.org/10.5072/FK2/TTRIBT>, Borealis, V1, UNF:6:QUEzj6HFqsjnX706WwMLdQ== [fileUNF]

< Hide Groups

English

Cross Tabulation Summary Statistics Download

Items per page: 25 1 - 25 of 65

Confirm Type	ID	Name	Label	Weight	View Summary Statistics	View Categories	View Questions	
Changes in Income	<input type="checkbox"/>	v6012811	PUMFID	Sequential Identification Number	v6012809			
Parents	<input type="checkbox"/>	v6012804	SRYP	Survey reference year	v6012809			
Benefits	<input type="checkbox"/>	v6012769	REGION6	Province or region				
Unemployment	<input type="checkbox"/>	v6012819	SEX	Sex of respondent				
Weight Variable	<input type="checkbox"/>	v6012801	AGECAT	Age of respondent - grouped				
Sociodemographics	<input type="checkbox"/>	v6012797	EDUC	Highest educational attainment				
Additional Sources of Funding	<input type="checkbox"/>	v6012779	TYPES	Survey type as determined for the EICS	v6012809			
Job Information	<input type="checkbox"/>	v6012822	WORKNOW	Respondent currently working at a job or business	v6012809			
Parental benefits for spouses	<input type="checkbox"/>							

Browse by Variable Groups

Select / Unselect Variables

View summary statistics, frequencies, question information

Summary Statistics

Employment Insurance Coverage Survey, 2021 [Canada] English

EICS-4428-E-2021_F1.tab

Labour Statistics Division, 2023, "Employment Insurance Coverage Survey, 2021 [Canada]", <https://doi.org/10.5072/FK2/TTRIBT>, Borealis, V1. UNF-6:QUEzi6HFasinX706WrMLdO== [fileUNF]

< Hide Groups

All Variables Search

Confirm Type	ID	Name
Changes in Income	<input type="checkbox"/> v6012811	PUMFID
Parents	<input type="checkbox"/> v6012804	SRYR
Benefits	<input type="checkbox"/> v6012769	REGION6
Unemployment	<input type="checkbox"/> v6012819	SEX
Weight Variable	<input type="checkbox"/> v6012801	AGECAT
Sociodemographics	<input type="checkbox"/> v6012797	EDUC
Additional Sources of Funding	<input type="checkbox"/> v6012779	TYPES
Job Information	<input checked="" type="checkbox"/> v6012822	WORKNOW
Parental benefits for spouses	<input checked="" type="checkbox"/> v6012775	PARENT
Unassigned	<input type="checkbox"/> v6012815	LNDIMM_G
Economic Family		Identifies if a person is a landed immigrant

WORKNOW: Respondent currently working at a job or business

Values	Categories	Count	Count Percentage(%)	Weighted Count	Weighted Count Percentage(%)
6	Valid skip	3,124	44.124	3,361,659.873	45.113
1	Yes	1,866	26.356	1,940,991.525	26.048
8	Refusal	0	0	0	0
2	No	2,074	29.294	2,129,783.673	28.582
7	Don't know	0	0	0	0
9	Not stated	16	0.226	19,147.644	0.257

Navigation: Cross Tabulation | **Summary Statistics** | Download

Items per page: 25 | 1 - 25 of 65

Weight	View Summary Statistics	View Categories	View Questions
v6012809			
v6012809			
v6012809			
v6012809			
v6012809			
v6012809			
v6012809			

Analyse data

Select Variables For Cross Tabulation

- VAP_15AR - Nb of days vaped e-liquid with nicotine - pas
- VAP_15BR - Nb of days vaped e-liquid without nicotine -
- VAP_15CR - Nb of days vaped e-liquid didn't know what it
- VAP_20R - Nb of times pick up vaping device - Day respo
- VAP_21R - Nb of puffs taken each time vaping device is
- VAP_30R - Flavour respondent vape most often
- VAP_40AR - Where get vaping devices - Vape shop
- VAP_40BR - Where get vaping devices - Convenience store/
- VAP_40CR - Where get vaping devices - Supermarket/grocer
- VAP_40DR - Where get vaping devices - Online



Rows
 VAP_35R - Main reason for vaping - currently



Columns
 AGEGROUP - Age group of person



Cross Tabulation Table

AGEGROUP - Age group of person	15 to 19 years old	20 to 24 years old	25 to 34 years old	35 to 44 years old	45 to 54 years old	55 to 64 years old	65 years old and older	Valid skip	Don't know	Refusal	Not stated	Total
VAP_35R - Main reason for vaping - currently												
By curiosity, you just wanted to try it	59 0.6%	45 0.45%	7 0.07%	3 0.03%	3 0.03%	1 0.01%	1 0.01%	0 0%	0 0%	0 0%	0 0%	119 1.2%
Because you enjoy it	72 0.73%	102 1.03%	9 0.09%	7 0.07%	2 0.02%	1 0.01%	1 0.01%	0 0%	0 0%	0 0%	0 0%	194 1.96%
To reduce stress or calm you down	87 0.88%	105 1.06%	7 0.07%	8 0.08%	2 0.02%	0 0%	3 0.03%	0 0%	0 0%	0 0%	0 0%	212 2.14%
To quit smoking cigarettes	6 0.06%	37 0.37%	10 0.1%	11 0.11%	12 0.12%	2 0.02%	6 0.06%	0 0%	0 0%	0 0%	0 0%	84 0.85%
To cut down on smoking cigarettes	1 0.01%	28 0.28%	4 0.04%	8 0.08%	3 0.03%	5 0.05%	4 0.04%	0 0%	0 0%	0 0%	0 0%	53 0.53%
To use when you cannot/are not allowed to smoke cigarettes	2 0.02%	8 0.08%	0 0%	2 0.02%	2 0.02%	1 0.01%	0 0%	0 0%	0 0%	0 0%	0 0%	15 0.15%
To avoid returning to smoking cigarettes	2 0.02%	41 0.41%	15 0.15%	8 0.08%	12 0.12%	3 0.03%	4 0.04%	0 0%	0 0%	0 0%	0 0%	85 0.86%
Other	31 0.31%	49 0.49%	3 0.03%	6 0.06%	2 0.02%	2 0.02%	2 0.02%	0 0%	0 0%	0 0%	0 0%	95 0.96%
	1775	1087	641	612	841	1100	1764	0	0	0	0	6070

Download and subset

Canadian Tobacco and Nicotine Survey, 2021 English

CTNS-5305-E-2021_F1.tab

Statistics Canada, 2023, "Canadian Tobacco and Nicotine Survey, 2021", <https://doi.org/10.5072/FK2/ZHT5KX>, Borealis, V1, UNF:6:15CjFBAeehXclm89HK6dpA== [fileUNF]

< Hide Groups Cross Tabulation Summary

Download

Download subset

Download original file

Download tab-delimited file

Download RData format file

Download variable metadata

Search

Items per page 25

	ID	Name	Label	Weight			
<input type="checkbox"/>							
<input checked="" type="checkbox"/>	v6423547	AGEGROUP	Age group of person				
<input checked="" type="checkbox"/>	v6423552	ALC_05	Frequency drank alcoholic beverage - past 30 days				
<input checked="" type="checkbox"/>	v6423564	ALC_10	Freq drank 4 drinks or more on one occasion - past 12 months				
<input checked="" type="checkbox"/>	v6423571	CAN_05AR	Smoked cannabis - lifetime				
<input checked="" type="checkbox"/>	v6423570	CAN_05BR	Age first time smoked cannabis				
<input type="checkbox"/>	v6423525	CAN_10AR	Frequency smoked cannabis - past 30 days				
<input type="checkbox"/>	v6423539	CAN_10BR	Nb of days smoked cannabis at least once past month - grouped				
<input type="checkbox"/>	v6423524	CAN_15AR	Frequency smoked cannabis/tobacco mix - past 30 days				
<input type="checkbox"/>	v6423569	CAN_15BR	Nb of days smoked can./toba. mix at least once a week - past 30 days				
<input type="checkbox"/>	v6423529	CAN_17R	Frequency consumed edibles - past 30 days				

Further Integrations

- Depositing and sharing tabular data in Dataverse
 - Improvements across tabular data ingest and Data Explorer (weighting cross-tabs, exporting charts, more visualizations, Data Documentation Initiative (DDI) support)
- Using Dataverse APIs you can build your own data exploration tools
- Data Explorer is an open, ready to use tool
- Data Explorer can be built into integrated data search sites (e.g.

Data catalogs

The screenshot shows the Odesi data catalog interface. At the top, there's a search bar with 'stress' entered and a dropdown menu set to 'Variable'. Below the search bar, there are filters for 'Collection' (All collections selected), 'Sort' (Relevance), and 'Date (YYYY)' (From and To). A blue 'SEARCH' button is visible. On the right side, there's a description of Odesi: 'Odesi is a Canadian social science data repository and online exploration and analysis tool. It contains 5,700+ datasets curated by academic libraries in Canada.' Below the search bar, the results are displayed as 'Results 1 to 20 of 513' and 'Page 1 of 26'. The first result is 'Enquête sociale générale 2010, Cycle 24 [Canada]: Bien-être et stress lié au manque de temps, Fichier principal'. The second result is 'General Social Survey, Cycle 24, 2010 [Canada]: Time-Stress and Well-Being, Main File [version 5]'. On the right side, there's a 'Limit your search:' section with a list of collections: Odesi 291, DU 235, Dlimt 184, POP 52, Borealis 38, and More. Below that, there's a 'Series' section with: Canadian Tobacco Use Monitoring Survey 40, Enquête De Surveillance De L'usage Du Tabac Au Canada 34, and Canadian Community Health Survey 24.

The User Experience: Permissions and Access

Permissions & Access: Flexible Permission System

- Supports multiple workflows by controlling who can deposit files on your behalf, and who can download files
- Roles are defined as a set of permissions to grant to users or to groups
- Groups can be defined statically or dynamically (e.g. users logging in from the same institution, via Shibboleth)

Access control:

- Custom terms
- Restricted data must clearly define terms of access
- “Request access” workflow

Users/Groups ⌵ All the users and groups that have access to restricted files in this dataset.

👤 Grant Access to Users/Groups

Include Deleted Files

0 Users/Groups

User/Group Name (Affiliation)	ID	Files	Access
There are no users or groups with access to the restricted files in this dataset.			

Restricted Files ⌵ All the restricted files in this dataset.

Include Deleted Files

0 Restricted Files

File Name	Published	Users/Groups	Access
There are no restricted files in this dataset.			

File Type: All ▾ Access: All ▾ ⌵ Sort

1 to 10 of 55 Files 📄 Download 🔒 Request Access

	boxplot_2d.py Python Source Code - 4.5 KB Published Jul 13, 2023 0 Downloads MD5: 093...1c9	
	converting_TSIS-1_to_UVspec_compatible_format.ipynb Jupyter Notebook - 72.8 KB Published Jul 13, 2023 0 Downloads MD5: 6be...0d1	
	convolution-1.py Python Source Code - 4.5 KB Published Jul 14, 2023 0 Downloads MD5: eb2...304	

Permissions & Access: Flexible Permission System

- Supports multiple workflows by controlling who can deposit files on your behalf, and who can download files
- Roles are defined as a set of permissions to grant to users or to groups
- Groups can be defined statically or dynamically (e.g. users logging in from the same institution, via Shibboleth)

Access control:

- Custom terms
- Restricted data must clearly define terms of access
- “Request access” workflow
- Public, Restricted, Embargoed

Access

Public (17,183)

Restricted (275)

Embargoed then Public (2)

Access: EmbargoedThenPublic ✕

1 to 2 of 2 Results

2-neu.txt **Draft** **Unpublished** **Embargoed**



Mar 16, 2023 - [gezippt](#)

Plain Text - 9 B - MD5: ca8...198 ↵

text text

7610-how-to-use-multilingual-external-module (1).docx **Embargoed**



Mar 16, 2023 - [Test dataset](#)

MS Word - 239.0 KB - MD5: 1bc...840 ↵

Permissions & Access: Flexible Permission System

- Supports multiple workflows by controlling who can deposit files on your behalf, and who can download files
- Roles are defined as a set of permissions to grant to users or to groups
- Groups can be defined statically or dynamically (e.g. users logging in from the same institution, via Shibboleth)

This dataset will be published under the terms specified. Please expect that proper credit is given via citation.

CC0 1.0

- CC BY-NC-SA 4.0
- CC BY-ND 4.0
- CC BY-SA 4.0
- PDDL-1.0
- ODC-By 1.0
- ODbL 1.0
- OGL UK 3.0
- Custom Dataset Terms

Restrict Access

Restricting limits access to published files. People who want to use the restricted files can request access by default. If you disable request access, you must add information about access to the Terms of Access field.

Learn about restricting files and dataset access in the [User Guide](#).

Request Access Enable access request

Terms of Access for Restricted Files

Access control:

- Custom terms
- Restricted data must clearly define terms of access
- “Request access” workflow
- Public, Restricted, Embargoed
- Licenses, Custom Licenses

Restricted Files + Terms of Access

Restricted Files ?	There is 1 restricted file in this dataset.
Request Access ?	Users may request access to files.

Permissions & Access: Flexible Permission System

- Supports multiple workflows by controlling who can deposit files on your behalf, and who can download files
- Roles are defined as a set of permissions to grant to users or to groups
- Groups can be defined statically or dynamically (e.g. users logging in from the same institution, via Shibboleth)



Restricted Files ^ All the restricted files in this dataset.

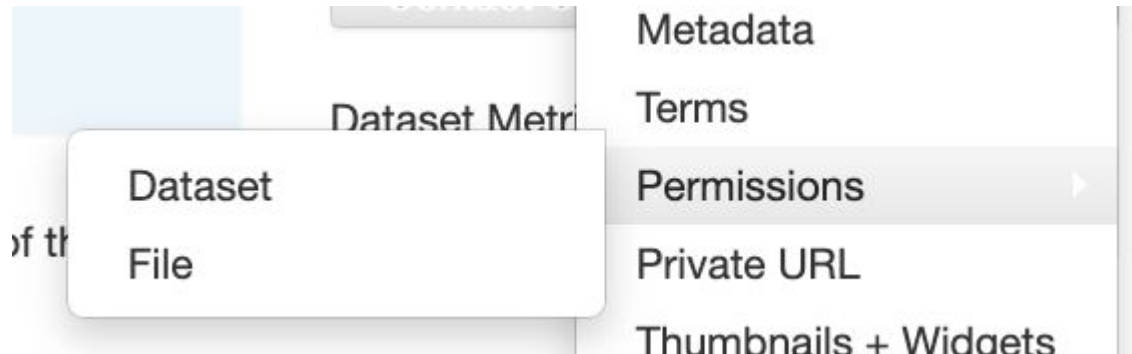
Include Deleted Files

1 Restricted File

File Name	Published	Users/Groups	Access
1643650411997.jpg	Restricted	0 Users/Groups	+ Assign Access

Access control:

- Custom terms
- Restricted data must clearly define terms of access
- “Request access” workflow
- Public, Restricted, Embargoed
- Licenses, Custom Licenses
- Dataset or File level download options



Permissions & Access: Flexible Permission System

- Supports multiple workflows by controlling who can deposit files on your behalf, and who can download files
- Roles are defined as a set of permissions to grant to users or to groups
- Groups can be defined statically or dynamically (e.g. users logging in from the same institution, via Shibboleth)

Access control:

- Custom terms
- Restricted data must clearly define terms of access
- “Request access” workflow
- Public, Restricted, Embargoed
- Licenses, Custom Licenses
- Dataset or File level download options

Prueba Demo Consorcio


Draft **Unpublished**

Saavedra, Paula, 2023, "Prueba De DRAFT VERSION" [?](#)

[Cite Dataset](#) [Learn about](#)

Description [?](#) pruebas 2

Subject [?](#) Agricultural Science

License/Data Use Agreement  CC0

Publish Dataset ▾

Edit Dataset ▾

Contact O

Dataset Metr

0 Download

- Files (Upload)
- Metadata
- Terms
- Permissions
- Private URL
- Thumbnails + Widgets

Delete Dataset



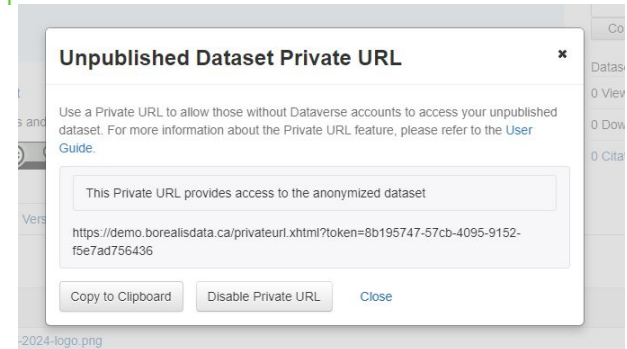
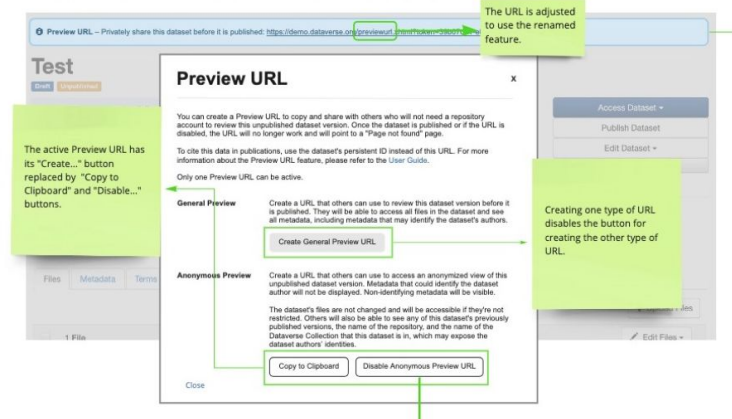
Permissions & Access: Flexible Permission System

- Supports multiple workflows by controlling who can deposit files on your behalf, and who can download files
- Roles are defined as a set of permissions to grant to users or to groups
- Groups can be defined statically or dynamically (e.g. users logging in from the same institution, via Shibboleth)

Access control:

- Custom terms
- Restricted data must clearly define terms of access
- "Request access" workflow
- Public, Restricted, Embargoed
- Licenses, Custom Licenses
- Dataset or File level download options

Depositor clicks on the "Create Anonymous Preview URL" button, the info banner with the URL appears, and the popup updates.

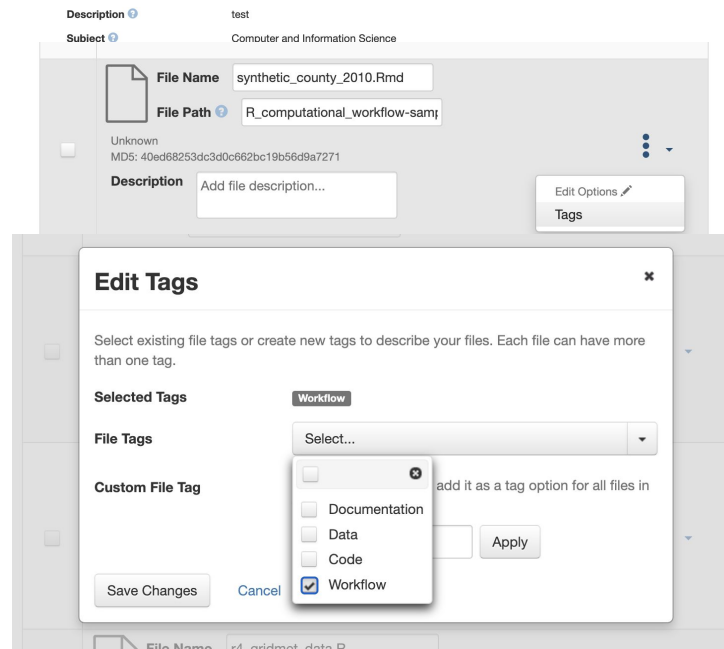


<https://github.com/IQSS/dataverse/issues/8185>

The User Experience: Data Exploration

Computational workflow and files support

- External tools exist to support **workflows** and **reproducibility**
- specific metadata to better support discoverability (e.g. the new “Dataset Feature” facet)
- Automatic checksum validation on **BagIt** file upload (based on the BagIt manifest)
- **Binder** support (to run Jupyter notebooks)
[soon a connection to NERC/MOC]



Dataverse & BagIt

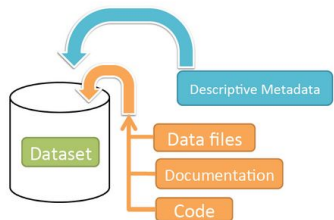
- Generates BagIt zip file with complete metadata & all datafiles for a Dataverse dataset
- Conforms to RDA recommendations & includes complete JSON-LD/RDF metadata using [OAI-ORE](#) structure
- Imports BagIt packages as datasets, providing round-trip, export/import capability.

Dataverse Archives Using BagIT

- [Texas Data Repository](#)
(Duracloud/Chronopolis)
- [Qualitative Data Repository](#)
(Google Cloud)
- [Harvard Data Commons](#)
Workflow developed for depositing archival bags into the Harvard DRS repo (not in production yet)

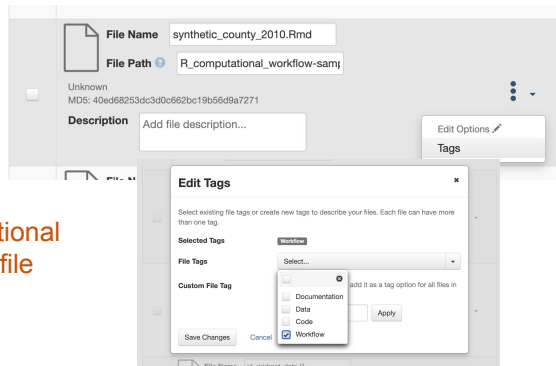
File Types, Format, Documentation

Schematic Diagram of a Dataset in Dataverse 4.0

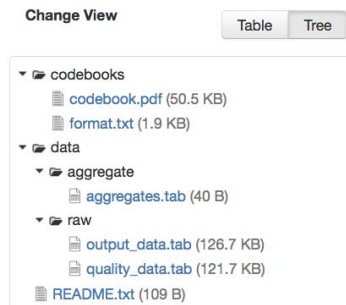


Container for your data, documentation, and code.

Computational workflow file support



Folder hierarchy support



Tabular data support

Supported File Formats

Tabular Data ingest supports the following file formats:

File format	Versions supported
SPSS (IPOR and SAV formats)	7 to 22
STATA	4 to 15
R	up to 3
Excel	XLSX only (XLS is NOT supported)
CSV (comma-separated values)	(limited support)

File level access control

Access

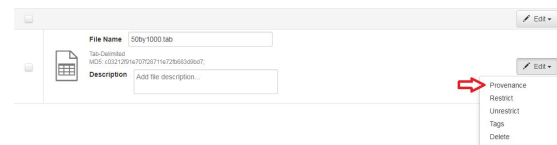
Public (1,784,339)

Restricted (50,591)

Embargoed then Public (154)

Embargoed then Restricted (14)

File level Provenance support



The User Experience: Search and Browse Functionality

Search:

- Keyword search
- Advanced search
- Funding agency
- Faceted search
- Sorting
- Cross repository integration

Keyword Term

- AFRICA (328)
- AFRICA SOUTH OF SAHARA (309)
- EAST AFRICA (162)
- ASIA (148)
- health (145)

Funding Information Agency

- United States Agency for International Development (USAID) (247)
- Bill and Melinda Gates Foundation (BMGF) (77)
- Bill and Melinda Gates Foundation (31)
- World Bank (24)
- Bill & Melinda Gates Foundation (BMGF) (23)



Dataset details:

- Detailed dataset page
- Metadata
- Description
- Authors
- Citation
- Download options
- Documentation, readme, code...
- Multiple format download options

Browse:

- Categories/Subjects
- Featured Datasets/collections

ModuleC_31member.tab
Data/
Tabular Data - 202.7 KB
Published Aug 2, 2023
2 Downloads
14 Variables, 245 Observations UNF:6:q2N...11A==

ModuleD_hh.tab
Data/
Tabular Data - 36.1 KB
Published Aug 2, 2023
0 Downloads
13 Variables, 1000 Observations UNF:6:Krb...10Q==

ModuleL5-5_var.tab
Tabular Data - 110.4 KB
Published Aug 2, 2023
0 Downloads
21 Variables, 1732 Observations UNF:6:MrOO...+qg==

SurveyInstrument.pdf
Adobe PDF - 1.1 MB
Published Aug 2, 2023
0 Downloads
MDS: cc7...e37

ModuleA_hh.tab
Data/
Tabular Data - 36.8 KB
Published Aug 2, 2023
0 Downloads
9 Variables, 1200 Observations UNF:6:brvA...1QA==

IFPRI Dataverse Home
(International Food Policy Research Institute)

Harvard Dataverse >

Contact Share

The International Food Policy Research Institute (IFPRI) views the products of its research, including research datasets, as global public goods, and is committed to enabling their widespread distribution and use.

This is in keeping with the IFPRI Research Data Management and Open Access (RDMOA) Policy and the CGIAR Open Access and Data Management Policy.

The IFPRI Dataverse comprises datasets collected during the course of IFPRI research. These datasets include tables in various standard formats, survey instruments, codebooks, metadata, and other associated documentation. The Terms of Use require proper attribution of these datasets to IFPRI and any named authors.

Please direct questions about IFPRI datasets to IFPRI-Data or IFPRI-Library. For information on the latest resources and news on research data management, please visit IFPRI's Research Data website.

ASTI led by IFPRI
Agricultural Science and Technology Indicators (ASTI) Dataverse

Nexus SAMs
Nexus SAM

HarvestChoice
BETTER CHECKS, BETTER LIVES
IFPRI HarvestChoice Dataverse

IMPACT
International Model for Policy Analysis of Agricultural Commodities and Trade (IMPACT) Dataverse

Search this dataverse... Advanced Search

Dataverses (8)
Datasets (791)
Files (13,127)

Dataverse Category
Research Project (7)
Research Group (1)

1 to 10 of 790 Results

Ethiopia Groundwater Games for Resource Governance **Engaged!**
Aug 2, 2023
International Food Policy Research Institute (IFPRI); Haramya University, 2023. "Ethiopia Groundwater Governance". <https://doi.org/10.21203/rs.3.rs-3127233>. Harvest Dataverse, 31

Name (A-Z)
Name (Z-A)
Newest
Oldest

Ask the Data



File Tools Open in New Window

Ask the Data I want to know

View Data

What do you want to know about?

Answer please

This data is about households and their members

Browse: Recently added

File Tools Metadata Viewers

Tell me what you want to know about this data about?

Answer please

This data is about households and their members.

HHID	memberID	cs1	cs2	cs3	cs4	cs5	cs6	cs7	cs8	cs9	cs10	cs11	cs12	cs13
101001	1	2	1	2	1	2	4	1	1	0	1	1	1	1
101001	2	1	2	26	4	2	8	4	1	0	1	1	1	1
101002	1	2	1	25	4	2	4	3	1	0	1	1	1	1
101002	2	1	2	27	4	2	4	3	1	0	1	1	1	1
101002	3	1	3	5										
101002	4	2	3	2										
101003	1	2	0	76	1	2	5	3	1	2	2	1	1	1
101003	2	2	4	27	1	2	5	3	1	2	2	1	1	1
101003	3	1	3	6										
101003	4	2	3	4										



Thank You!

The Dataverse Project

<https://dataverse.org/>

The Harvard Dataverse repository

<https://dataverse.harvard.edu/>

Integrating Dataverse and DSpace

<https://osf.io/72w4m>

The Dataverse Guide

<https://guides.dataverse.org/en/latest/>

Dataverse on Github

<https://github.com/IOSS/dataverse>

Dataverse google community

<https://groups.google.com/g/dataverse-community>

References:

Dataverse - [Goals, Roadmaps, and Releases](#)

Introduction to [Metadata in Dataverse Repositories](#): For
Researchers and Support Staff

Introduction to [Data Handling on the Dataverse](#) Platform

Dataverse Software - [New Features and Future Plans](#)