



# The Dataverse Project UbuntuNet-Connect, 2024

Sonia Barbosa

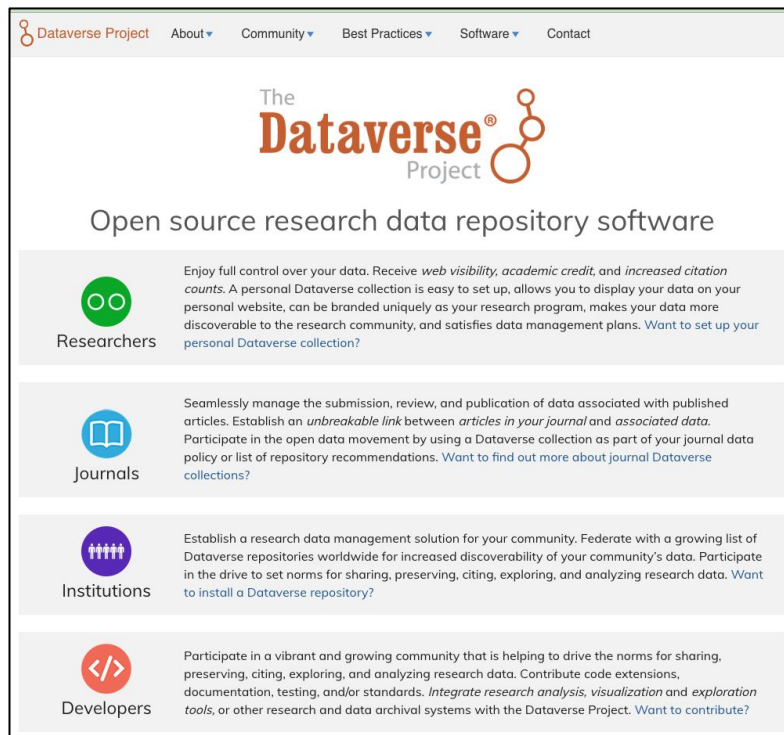
# Agenda

1. Dataverse Project History and Timeline(s)
2. Why the Dataverse Project and Use Cases
3. Dataverse Tools and Features
4. Computing on the Data
5. Grants and Collaborations
6. Global Dataverse Community
7. Dataverse Project Today & Tomorrow
8. Collaborate with Dataverse

# Dataverse Project

## About the Dataverse Project

- An **open-source** repository to publish, cite, and archive **research data**
- Built to support multiple types of **data, users, and workflows**
- Developed at Harvard's Institute for Quantitative Social Science (IQSS)
- Established in 1997, Dataverse Network Launch in 2006
- Development funded by IQSS and with grants, in collaboration with institutions around the world
- 2022 - present: Re-Architecture
- Core team
  - @ IQSS - developers, UX/UI, metadata specialists, curation team, leadership team
  - Key contributors from the community with full privileges as IQSS team



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

- Researchers** (green icon): 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** (blue icon): 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** (purple icon): 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** (red icon): 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 Project Goals

- Grow the Dataverse **community**
- Empower the open source community to explore and implement new **Dataverse applications, tools, and services**
- Develop the capability to handle **sensitive data and big data**
- Expand **data and metadata** features for existing and new disciplines
- Expand **archival and preservation** features
- Increase **interoperability** through the implementation of **standards**
- Increase contributions from the **open source development community**
- Improve the **Dataverse user experience**

King, G. (2007). An Introduction to the Dataverse Network as an infrastructure for data sharing. *Sociological Methods & Research*, 36(2), 173–199.

Altman & King (2007). A Proposed Standard for the Scholarly Citation of Quantitative Data, *D-Lib Magazine*, 13, 3/4(March/April, 2007).

Wilkinson, et al. (2016). The FAIR guiding principles for scientific data management and stewardship. *Scientific Data*, 3.

**1997**

Project Begins  
Focus on preserving & sharing social sciences data  
Harvard-MIT Data Center collaboration  
Murray Data Archive

**2006**

The Dataverse Network 1.0  
Harvard Dataverse Repository established


**2007**

 Announcing the Dataverse Network

**2008**

Odum Archive becomes first Dataverse outside Harvard

**2015**

Data Citations  
 First Open Dataverse Community Meeting  
Dataverse 4.0

**2016**

FAIR data

**2023**

 100th Dataverse installation  
Dataverse 6.0  
Dataverse Re-architecture  
 AI comes to Harvard Dataverse

**2024**

 The Next Frontier!  
Large data  
Computing on data

Altman, Borgman, Crosas & Matone. (2015). An introduction to the joint principles for data citation. *Bulletin of the Association for Information Science and Technology*, 41(3), 43–45.

# What Challenges does The Dataverse Project Solve?

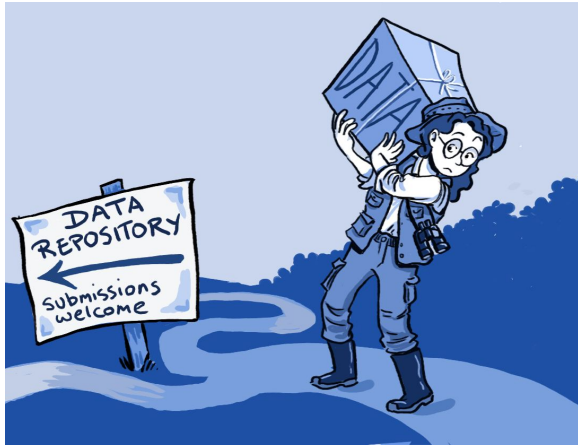


Illustration credit: Ainsley Seago.  
[doi:10.1371/journal.pbio.1001779.g001](https://doi.org/10.1371/journal.pbio.1001779.g001)

- Dataverse helps researchers manage the full lifecycle of data—from creation and analysis to publication and preservation—while meeting standards for openness and reproducibility

## Challenges

- Data inaccessibility
- Lack of data credit
- Data Preservation concerns
- Complex data management

## Solutions

- **Centralized repository**
- **Data Sharing**
- **Data citation system**
- **Credit for Data**
- **Data Preservation**
- **Data Discoverability**
- **Versioning**
- **Data Security**

# DATVERSE REPOSITORIES - A WORLD VIEW

123 Installations




  
DataverseNO



  
HARVARD  
Dataverse

 Dataverse Search - User Guide - Support

MELDATA

 Dataverse

An Institute of  
NANYANG TECHNOLOGICAL UNIVERSITY  
SINGAPORE

 World Agroforestry (ICRAF)

# Use Cases

**HARVARD**  
Dataverse

[Add Data](#) ▾ [Search](#) ▾ [About](#) [User Guide](#) [Support](#) [Sign Up](#) [Log In](#)

**Deposit and share your data. Get academic credit.**

Harvard Dataverse is a repository for research data. Deposit data and code here.

[Add a dataset +](#)

**Organize datasets and gather metrics in your own repository.**

A dataverse is a container for all your datasets, files, and metadata.

[Add a dataverse +](#)

**Publishing your data is easy on Harvard Dataverse!**

Learn about getting started creating your own dataverse repository here.

[Getting started ↗](#)

---

Find data across research fields, preview metadata, and download files

Search over 182,900 datasets...  [VIEW ALL DATA >](#)

**Featured** **COVID-19 Data Collection**

A curated collection of COVID-19 data deposited in the Harvard Dataverse repository.

[Dataverses \(6,831\)](#)

[Datasets \(182,901\)](#)

[Files \(3,409,257\)](#)

**Dataverse Category**

- [Research Project \(2,556\)](#)
- [Researcher \(2,137\)](#)
- [Organization or Institution \(540\)](#)
- [Research Group \(478\)](#)
- [Journal \(147\)](#)
- [Laboratory \(126\)](#)
- [Department \(78\)](#)
- [Teaching Course \(45\)](#)

**Publication Year**

2024	(476)
2023	(271)
2022	(820)
2021	(1,047)
2020	(842)
2019	(492)
2018	(542)
2017	(509)
2016	(521)
2015	(490)
2014	(234)
2013	(124)
2012	(106)
2011	(97)
2010	(55)
2009	(42)
2008	(49)
2007	(114)

**Subject**

- [Social Sciences \(63,535\)](#)
- [Arts and Humanities \(36,538\)](#)
- [Medicine, Health and Life Sciences \(9,917\)](#)
- [Earth and Environmental Sciences \(9,102\)](#)
- [Law \(5,747\)](#)
- [Agricultural Sciences \(4,540\)](#)
- [Computer and Information Science \(3,901\)](#)
- [Other \(2,675\)](#)
- [Business and Management \(2,444\)](#)
- [Engineering \(2,105\)](#)
- [Physics \(1,674\)](#)
- [Astronomy and Astrophysics \(1,151\)](#)
- [Chemistry \(911\)](#)
- [Mathematical Sciences \(620\)](#)

Covid19 data sharing!

**HARVARD**  
Dataverse

**ILRI**  
International Livestock Research Institute - ILRI Dataverse

Harvard Dataverse >

**HARVARD**  
Dataverse

**PSI**  
Population Services International (PSI) Dataverse

Harvard Dataverse >

**HARVARD**  
Dataverse

**IFPRI**  
International Food Policy Research Institute

Harvard Dataverse >

**HARVARD**  
Dataverse

**CHOP**  
Children's Hospital of Philadelphia RESEARCH INSTITUTE

Children's Hospital of Philadelphia Dataverse

Harvard Dataverse >

**HARVARD**  
Dataverse

**ASTI**  
led by IFPRI

Harvard Dataverse >

**HARVARD**  
Dataverse

**Nexus SAMs**

Harvard Dataverse >



**borealis** The Canadian Dataverse Repository  
Le dépôt Dataverse canadien

**Store, share, publish  
and discover research  
data!**

EXPLORE BOREALIS

ABOUT US

Search Borealis



## Eastern Canada

Acadia University  
Cape Breton University  
Dalhousie University  
Memorial University  
Mount Allison University  
Mount Saint Vincent University  
St. Francis Xavier University  
Saint Mary's University

## Ontario

Algoma University  
Brock University  
Carleton University  
Durham College  
Fanshawe College  
Lakehead University  
Laurentian University  
McMaster University  
Nipissing University  
OCAD University  
Ontario Tech University  
Queen's University  
Royal Military College  
Toronto Metropolitan University  
Trent University  
University of Guelph  
University of Ottawa  
University of Toronto  
University of Waterloo  
University of Windsor  
Western University  
Wilfrid Laurier University  
York University

## Quebec

Bishop's University  
Concordia University  
École de technologie supérieure  
École nationale d'administration publique  
Fédération des cégeps  
HEC Montréal  
Institut national de la recherche scientifique  
McGill University  
Polytechnique Montréal  
Université Laval  
Université de Montréal  
Université du Québec à Chicoutimi  
Université du Québec à Montréal  
Université du Québec à Rimouski  
Université du Québec à Trois-Rivières  
Université du Québec en Abitibi-Témiscamingue  
Université du Québec en Outaouais  
Université de Sherbrooke  
Université TÉLUQ

## Western Canada

Athabasca University  
Brandon University  
Concordia University of Edmonton  
MacEwan University  
Mount Royal University  
Royal Roads University  
Thompson Rivers University  
Trinity Western University  
University of Alberta  
University of British Columbia  
University of Calgary  
University of Lethbridge  
University of Manitoba  
University of Northern British Columbia  
University of Regina  
Université de Saint-Boniface  
University of Victoria  
University of Winnipeg  
Vancouver Island University



**borealis**



# Repository of open research data of the National Academy of Sciences of Ukraine

Search datasets in DataverseUA

Search

Go to the repository

Data placement request

Recently published datasets

## Advantages

DataverseUA is a public platform of the Open Data Repository, available to researchers of the institutes of the National Academy of Sciences of Ukraine for the purpose of hosting, sharing and reusing research data. This helps to ensure the multiple use of research data, its availability, the ability to interact with different types of data and the implementation of operational search (FAIR principles).

DataverseUA supports creating custom terms of use and restrictions to control access to your research data. DataverseUA provides guaranteed access to data sets, permanent identifiers and special conditions for saving research data, creating backup copies for long-term storage.

Will post research data on DataverseUA, making it more accessible to the research community on the Internet. You can count on the Dataverse platform to work well with all types of data and provide access to advanced search capabilities, allowing your researchers to find the data they need and extract valuable insights from the data. The Dataverse platform is secure, interoperable, scalable and

<https://opendata.nas.gov.ua/>

DataverseUA

Metrics 2 Downloads

Contact Share

Data repository of the institutes of the National Academy of Sciences of Ukraine.



Search this dataverse...

Advanced Search

Datasets (4)

Datasets (0)

Files (0)

**Dataverse Category**

Organization or Institution (2)

Department (1)

Research Group (1)

**Publication Year**

2023 (2)

2022 (2)

**Subject**

Chemistry (2)

Physics (2)

1 to 4 of 4 Results

Sort

UHV ANALYSIS SYSTEM Center (Frantsevych Institute for Problems of Materials Science of NAS of Ukraine)

Sep 19, 2023 Frantsevych Institute for Problems of Materials Science of NAS of Ukraine

Center for collective use of appliances "High-vacuum analytical system UHV-ANALYSIS-SYSTEM" is designed to conduct research in the field of electronic structure, elemental and phase composition of the surface of solids without their destruction. This center was created on the basis...

SPM&RS Center (Kurdyumov Institute for Metal Physics of the NAS of Ukraine)

Sep 19, 2023 G.V. Kurdyumov Institute for Metal Physics of the NAS of Ukraine

The collective use center "Centre of scanning probe microscopy and resonance spectroscopy" (SPM&RS-Centre) was created in 2012 in the structure of the GV Kurdyumov Institute for Metal Physics National Academy of Sciences of Ukraine for the most rational use of unique equipment...

Frantsevych Institute for Problems of Materials Science of NAS of Ukraine (National Academy of Sciences of Ukraine)

February 9, 2022

Frantsevych Institute for Problems of Materials Science of NAS of Ukraine is a leading center of scientific and technical developments in the field of theoretical foundations of new materials formation, technology for production and manufacture of products from these materials w...

GV Kurdyumov Institute for Metal Physics of the NAS of Ukraine (National Academy of Sciences of Ukraine)

January 7, 2022

GV Kurdyumov Institute for Metal Physics of the NAS of Ukraine (hereinafter, Institute) was founded in 1945. It is one of the largest scientific centers of fundamental research in the field of metal physics in Ukraine and Europe. The scientists of the Institute carry out re...

JPL Open Repository

(JPL)

Root >

Contact Share

The JPL Open Repository (JOR) replaced the JPL Technical Report Server (TRS) in 2023. It is a repository for digital copies of technical publications authored by JPL employees. It includes journal articles (the Final Accepted Version), meeting papers, presentations, and other publications cleared for external distribution from 1992 to the present. It also includes research datasets from 2022 on.

Years covered: 1992 to the present.

- 1992-2021 records were originally on the JPL Technical Report Server (TRS) on Dspace system; those records were migrated to Dataverse in 2022.
- Records with publication date 2022 or later are native to the Dataverse.
- Data sets: 2022 to present

Read full Description [+]

Search this dataverse...   [Advanced Search](#)

**Dataverses (0)**

**Datasets (44,022)**

**Files (44,868)**

**Author**

- Crisp, David (311)
- Bar-Cohen, Y. (261)
- Chien, Steve (186)
- Eldering, Annmarie (142)
- Bar-Cohen, Yoseph (141)

More...

**Document Type**

- Other (18,184)
- Slides (Viewgraph) (15,177)
- Meeting Paper (7,049)
- Journal Article (3,154)
- Report (413)

More...

**Keywords**

Earth (4)

More (4)

1 to 10 of 44,022 Results

Sort

**Deep Space Optical Communications Technology Demonstration**

Oct 20, 2024

Abhijit Biswas, 2024, "Deep Space Optical Communications Technology Demonstration", <https://doi.org/10.48577/jpl.BTYYYV>, Optica Laser Congress 2024, JPL Open Repository

Space-to-ground high data-rate (6.25 to 267 Mb/s) deep space optical communications over distances of 0.2 to 2.5 astronomical units (AU) were demonstrated with a space payload hosted by the Psyche Mission spacecraft, (launched, October 13, 2023) and ground assets. The system impl...

**Revealing the Lifecycle of Interstellar Dust with Polarimetry**

Oct 18, 2024

Brandon Hensley, 2024, "Revealing the Lifecycle of Interstellar Dust with Polarimetry", <https://doi.org/10.48577/jpl.C7C5GL>, Observatoire astronomique de Strasbourg Colloquium, JPL Open Repository

No abstract available.

**Autonomous robotics and AI are driving NASA's progress in planetary exploration**

Oct 18, 2024

Vandi Verma, 2024, "Autonomous robotics and AI are driving NASA's progress in planetary exploration", <https://doi.org/10.48577/jpl.M4P8IA>, 2024 United Nations AI for Good Summit, JPL Open Repository

No abstract available.

**Open Distillation Characterization of Evanescent Absorption: Observation, Theory, and Experiment**

Root > JPL Open Repository > Deep Space Optical Communications Technology Demonstration >

# CL24\_3352.pdf

This file is part of "Deep Space Optical Communications Technology Demonstration".

Version 2.1

File Citation

Abhijit Biswas, 2024, "Deep Space Optical Communications Technology Demonstration", <https://doi.org/10.48577/jpl.BTYYYV>, Optica Laser Congress 2024, JPL Open Repository, CL24\_3352.pdf [fileName]

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

Dataset Citation


Abhijit Biswas, 2024, "Deep Space Optical Communications Technology Demonstration", <https://doi.org/10.48577/jpl.BTYYYV>, Optica Laser Congress 2024, JPL Open Repository

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

Preview Metadata Versions

Explore on Read Document

Conference title, upper and lower case, bolded, 1... 1 / 3 - 99% + [Print] [Refresh]



1

## Deep Space Optical Communications Technology Demonstration

**A. Biswas, M. Srinivasan, J. Allmaras, E. Alerstam, A. Velasco, E. Wollman**  
*Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109*  
Abhijit.Biswas@jpl.nasa.gov

**Abstract:** Space-to-ground high data-rate (6.25 to 267 Mb/s) deep space optical communications over distances of 0.2 to 2.5 astronomical units (AU) were demonstrated with a space payload hosted by the Psyche Mission spacecraft, (launched, October 13, 2023) and ground assets. The system implementation and results are summarized.

© 2019 California Institute of Technology. Government sponsorship acknowledged.

**1. Introduction**


*1.1. Motivation*

The demand for larger data volumes with higher data-rates, from NASA's deep space missions continues to grow [1]. Higher resolution science from around the solar system and human exploration from planets, will need streaming high-definition imagery and other high-rate data products. Existing robust state-of-art telecommunication systems are bandwidth

**HARVARD**  
Dataverse

Add Data Search About User Guide Support Sign Up Log In

**Politics of Genomics Dataverse** Home Page  
(Harvard University)



Harvard Dataverse >

The genomics revolution is underway. The effects of the companies' or employers' use of personal genetic information and DNA testing to discover one's genetic heritage (or public, policy-makers, and scholars have limited information on how the genomics revolution is affecting people's lives

Search this dataverse...

Datasets (0)

**Datasets (1)**

Files (3)

**Publication Year**  
2011 (1)


**Author Name**  
Jennifer Hochschild (1)  
Maya Sen (1)

**Author Affiliation**  
Harvard University (1)

**HARVARD**  
Dataverse

Add Data Search About User Guide Support Sign Up Log In

**COVID-19 Data Collection**



Harvard Dataverse >

This is a general collection of COVID-19 data deposited in the Harvard Dataverse (Harvard Library). Researchers who deposit their related content are encouraged to use the **contact** link if you have any questions about this collection.

Search this dataverse...  **Advanced Search**

**Datasets (134)**

Files (55,973)

**Dataverse Category**  
Research Project (78)  
Research Group (25)  
Researcher (18)  
Organization or Institution (4)  
Department (3)


**Metadata Source**  
Harvard Dataverse (2,362)  
Harvested (316)

**Publication Year**  
2024 (131)  
2023 (312)  
2022 (451)  
2021 (569)  
2020 (422)

**License**  
CC0 1.0 (1,420)  
Custom Terms (916)

**HARVARD**  
Dataverse

Add Data Search About User Guide Support Sign Up Log In

  
**Daniel Aldrich Dataverse** Northeastern University  
(Northeastern University)

Harvard Dataverse >

Search this dataverse...  **Advanced Search**

**Datasets (0)**

**Datasets (10)**

Files (27)

**Publication Year**  
2022 (1)  
2020 (3)  
2019 (2)  
2016 (1)  
2012 (1)


**Subject**  
Social Sciences (7)  
Medicine, Health and Life Sciences (2)  
Business and Management (1)  
Earth and Environmental Sciences (1)  
Engineering (1)

**Author Name**  
Aldrich, Daniel (7)  
Daniel P. Aldrich (2)  
Aldrich, Daniel P (1)  
Balcioglu, Zeynep (1)  
Berndt, Eric (1)

**Author Affiliation**  
Northeastern University (7)

**HARVARD**  
Dataverse

Add Data Search About User Guide Support Sign Up Log In

 **Brain Genomics Superstruct Project**

**Brain Genomics Superstruct Project (GSP) Dataverse** Home

(Harvard University)

Harvard Dataverse >

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.**

Search this dataverse...  **Advanced Search**

**Datasets (0)**

**Datasets (1)**


Files (15)

**Publication Year**  
2014 (1)

**Subject**  
Other (1)

**1 to 1 of 1 Result**

**Brain Genomics Superstruct Project (GSP)**  
May 15, 2015

 Buckner, Randy L.; Roffman, Joshua L.; Smoller, Jordan W., 2014, "Brain Genomics Superstruct Project (GSP)", <https://doi.org/10.7910/DVN/25853>, 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...

**American Journal of Political Science (AJPS) Dataverse** [ajps.org](#)

(Midwest Political Science Association)

[Harvard Dataverse >](#)

The *American Journal of Political Science* is committed to significant advances in knowledge and political science research. To find out more about our data integrity policies, please visit our website.

Search this dataverse...  [Advanced Search](#)

**Dataverses (0)**  
 **Datasets (722)**  
 **Files (17,915)**

**Publication Year**

[2024 \(47\)](#)  
[2023 \(65\)](#)  
[2022 \(49\)](#)  
[2021 \(73\)](#)  
[2020 \(70\)](#)

**License**

[Custom Terms \(539\)](#)  
[CC0 1.0 \(183\)](#)

**Author Name**

[Broockman, David \(9\)](#)  
[Hainmueller, Jens \(6\)](#)  
[Grossman, Guy \(5\)](#)  
[Imai, Kosuke \(5\)](#)  
[Mason, Andrew \(4\)](#)  
[Mason, Andrew \(4\)](#)

**1 to 10 of 722 Results**

**Replication Data for: The Financialization of Housing and Its Political Consequences**

Oct 16, 2024

Dancygier, Rafaela; Wiedemann, Andreas. "Replication Data for: The Financialization of Housing and Its Political Consequences", <https://doi.org/10.7910/DVN/PYDJKM>, Harvard Dataverse, V1

Institutional investors in residential real estate have become targets of political backlash against affordable housing. We argue that this backlash is not only about economic issues such as rising rents; it reflects a fundamental rejection of "financialized capitalism" that turns housing from a basic need into a speculative asset. Using novel geo-coded real estate transaction data, we document the extent of housing financialization cross-nationally and over time, and demonstrate that neighborhood-level exposure to financialization alone is insufficient to explain the widespread support to expropriate corporate landlords in a historic 2021 Berlin referendum. We then develop nationally-representative surveys to show that German citizens conceptualize housing as a social right and hold the state responsible for its under-provision. We demonstrate experimentally that arguments about housing financialization significantly raise support for expropriation beyond rent effects. Our findings suggest that financialized capitalism can unite diverse groups of voters in favor of housing socialism.

**Replication Data for: Can Interest Groups Influence the Electoral Prospects of Unions?**

Oct 7, 2024

Foulniaux, Alexander. 2024. "Replication Data for: Can Interest Groups Influence the Electoral Prospects of Unions?" <https://doi.org/10.7910/DVN/9KJGK>, Harvard Dataverse, V1

Unions sponsor electoral candidates around the world. We investigate how the electoral prospect of British Labour Party candidates is affected by the electoral prospect of British Labour Party candidates.

**Economic Risk Perceptions and Willingness to Underprivileged Groups in Vietnam**

Oct 7, 2024

Malesky, Edmund; Bonifal, Niccolo; Ruffalo, Joseph. 2024. "Economic Risk Perceptions and Willingness to Underprivileged Groups in Vietnam." <https://doi.org/10.7910/DVN/9KJGK>, Harvard Dataverse, V1

Economic risk perceptions and willingness to underprivileged groups in Vietnam. Field Experiment with Migrants and Other Underprivileged Groups in Vietnam.

Harvard Dataverse > American Journal of Political Science (AJPS) Dataverse >

### Replication Data for: The Financialization of Housing and Its Political Consequences

Version 1.0

Dancygier, Rafaela; Wiedemann, Andreas, 2024. "Replication Data for: The Financialization of Housing and Its Political Consequences", <https://doi.org/10.7910/DVN/PYDJKM>, Harvard Dataverse, V1

[Access Dataset](#)

Contact Owner Share

[Dataset Metrics](#)

21 Downloads

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

**Description**

Institutional investors in residential real estate have become targets of political backlash against affordable housing. We argue that this backlash is not only about economic issues such as rising rents; it reflects a fundamental rejection of "financialized capitalism" that turns housing from a basic need into a speculative asset. Using novel geo-coded real estate transaction data, we document the extent of housing financialization cross-nationally and over time, and demonstrate that neighborhood-level exposure to financialization alone is insufficient to explain the widespread support to expropriate corporate landlords in a historic 2021 Berlin referendum. We then develop nationally-representative surveys to show that German citizens conceptualize housing as a social right and hold the state responsible for its under-provision. We demonstrate experimentally that arguments about housing financialization significantly raise support for expropriation beyond rent effects. Our findings suggest that financialized capitalism can unite diverse groups of voters in favor of housing socialism.

**Subject**

Social Sciences

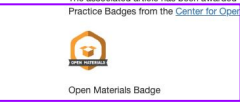
**Keyword**

Housing; Financialization; Socialism; Expropriation; Public Opinion; Social Movements; Germany

**Notes**

This dataset underwent an independent verification process, complying with the AJPS Verification Policy updated June 2023, which replicated the tables and figures in the primary article. For the supplementary materials, verification was performed solely for the successful execution of the replication process. The verification process was carried out by the Cornell Center for Social Sciences at Cornell University.

The associated article has been awarded the Open Materials Badge. Learn more about the Open Materials Badge from the [Center for Open Science](#).



Open Materials Badge

Proprietary data used in this study are not included in the replication package.

**License/Data Use Agreement**

[Custom Dataset Terms](#)

[Files](#) [Metadata](#) [Terms](#) [Versions](#)

- Reproducibility verification
- Submit for review workflow
- Open access badges

### AJPS Verification Policy

The corresponding author of a manuscript that is accepted for publication in the *American Journal of Political Science* must provide materials that are sufficient for other researchers to verify all of the analytic results that are reported in the manuscript. The document titled "American Journal of Political Science Preparing Verification Files" provides useful information about what is needed and how it should be organized. All verification files must be made available within the AJPS Dataverse, on the Harvard Dataverse Network. Do not make their verification files available elsewhere (e.g., their personal websites, repositories, etc.) as long as all of the necessary files are included in the AJPS Dataverse.

The corresponding author should prepare and upload verification materials to the AJPS Dataverse before submitting the final draft of the accepted manuscript. The *American Journal of Political Science Quick Reference for Uploading Verification Materials* provides information about creating a Dataset on the AJPS Dataverse. The *American Journal of Political Science Quantitative Data Checklist* and *American Journal of Political Science Qualitative Data Checklist* are designed to help authors ensure that they have provided all necessary materials.

When the final draft of the manuscript is submitted, the materials must be made available to the AJPS Dataverse. The materials confirm that they do, in fact, reproduce the analytic results reported in the article. Publication in the *American Journal of Political Science* is contingent on complete verification materials and successful verification of the materials.

AJPS articles that do not contain empirical data analyses are not subject to the verification policy. In some limited circumstances, an author may be granted an exemption from the verification policy. This exemption would allow the author to make some or all of the data used in an analysis. All other data (e.g., software commands, etc.) still must be provided. The primary author's request for an exemption is restricted to datasets and human subjects data. If an exemption is granted, then the author must include a note at the end of the published article explicating acknowledging the limitations on data availability. The request for an exemption must be approved explicitly by the AJPS Editor-in-Chief who retains final authority to decide whether the requested exemption is granted.

\*NOTE: Updates to the verification policy documents coming soon!

**License/Data Use Agreement** [Custom Dataset Terms](#)

[Files](#) [Metadata](#) [Terms](#) [Versions](#)

**Change View** [Table](#) [Tree](#)

- replication\_files
  - codebook.pdf (111.4 KB)
  - readme.txt (4.7 KB)
  - data
  - figures

## 2013 Photographs

(Harvard Museum of the Ancient Near East)

Harvard Dataverse > Leon Levy Expedition to Ashkelon Dataverse Collection > Photographs > 2013 Photographs >

# A13\_26171.jpg

Version 1.0



Master, Daniel M.; Stager, Lawrence E., 2022, "A13\_26171.jpg", <https://doi.org/10.7910/DVN/EBTORV>,  
Harvard Dataverse, V1

Cite Dataset ▾

Learn about [Data Citation Standards](#).

Access Dataset ▾

Contact Owner

Share

Dataset Metrics ⓘ

3 Downloads ⓘ

### Description ⓘ

Link to OCHRE database: <http://pi.lib.uchicago.edu/1001/org/ochre/69bbd137-0161-4f23-85f8-f78f74753cfa> (2022-03-01)

### Subject ⓘ

Arts and Humanities

### Keyword ⓘ

Archaeology

### License/Data Use Agreement

[Custom Dataset Terms](#)

Files

Metadata

Terms

Versions

1 File



A13\_26171.jpg

JPEG Image - 2.1 MB

Published Apr 19, 2022

3 Downloads

MD5: 415...ed3

<http://pi.lib.uchicago.edu/1001/org/ochre/69bbd137-0161-4f23-85f8-f78f74753cfa>

Photo



Standard Dataset Citation and DOI


Metrics

Linking to related projects

File page  
Metadata page  
Terms  
Versions

Tools/interoperability

**HARVARD Dataserve** | [Add Data](#) | [Search](#) | [About](#) | [User Guide](#) | [Support](#) | [Sign Up](#) | [Log In](#)



**Slavery, Abolition, Emancipation, and Freedom Collection** | Discover Digitized Primary Sources Detailing Black Experiences with Slavery, Abolition, and Freedom | (Harvard University)

Harvard Dataserve > Houghton Library Dataserve Collection >

**Project Background**  
 Houghton Library, Harvard University's largest rare books and manuscripts repository, is home to hundreds of thousands of materials relating to history across all collections related to Black history ranging from the 18th century through today, but have historically been difficult to discover amongst all the other materials. This together a curated collection of materials ranging from the Early Republic through Reconstruction.

In the summer of 2020, under the leadership of Digital Collections Program Manager Dorothy Berry, Houghton Library put a year-long pause on its digitization of curating a digital collection of materials relating to African American history and culture. Houghton's collections span from the earliest written materials on paper to 21st century book arts, and while our digitization has on the ground, we have not historically centered Black history. This project is a first foray into actively shifting the lens to Black experiences from the 18th through early 20th century.

[Read full Description \(+\)](#)

Search this dataserve... [Advanced Search](#)

**Dataserves (6)**  
**Datasets (1,229)**  
 Files (78,567)

**Publication Year**  
 2023 (1,229)

**Subject**  
 Arts and Humanities (1,228)  
 Social Sciences (1)

**Author Name**  
 Houghton Library (1,229)

**Author Affiliation**  
 Harvard University (1,229)

**Keyword Term**  
 Abolitionist Apologetics (1)  
 Abolitionist politics (1)  
 Abolitionists (1)  
 African American books (1)

**1 to 10 of 1,229 Results**


hou00124c01193	Feb 17, 2023	Houghton Library, 2023. "hou00124c01193". <a href="https://doi.org/10.7910/DVN/RFT3ME">https://doi.org/10.7910/DVN/RFT3ME</a> . Harvard Dataserve, V1. UNF:6:CSV656rH4yq3NmkZVbe= [file:UNF]
hou00124c01199	Feb 17, 2023	Higginson, Thomas Wentworth, 1823-1911. Post-war correspondence, 1866-1875.
hou00124c01199	Feb 17, 2023	Houghton Library, 2023. "hou00124c01199". <a href="https://doi.org/10.7910/DVN/PLT029">https://doi.org/10.7910/DVN/PLT029</a> . Harvard Dataserve, V1. UNF:6:G2q3b5YnczwwHd5WQew= [file:UNF]
hou0001c00078	Feb 17, 2023	United States. Army. South Carolina Volunteers, 1st. Correspondence, miscellaneous, 1863-1866.
hou0001c00078	Feb 17, 2023	Houghton Library, 2023. "hou0001c00078". <a href="https://doi.org/10.7910/DVN/9A9V6">https://doi.org/10.7910/DVN/9A9V6</a> . Harvard Dataserve, V1. UNF:6:G2q3b5YnczwwHd5WQew= [file:UNF]
hou0001c00078	Feb 17, 2023	Owen, Robert Dale, 1801-1877. Memo to S.G. (Samuel Gridley) Howe, New York, New York, 1863 July 17.

Files | **Metadata** | Terms | Versions

Search this dataset...

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

1 to 10 of 58 Files


 **hou00124c01193.xml**  
 XML - 442.2 KB  
 Published Feb 17, 2023  
 13 Downloads  
 MDS: a88..618

File associated with: Higginson, Thomas Wentworth, 1823-1911. Post-war correspondence, 1866-1875.  
 Origin of source: <https://nrs.harvard.edu/URN-3:FHCL.HOUGH:101293733>

City:Havana Country:United States Created:1866 Created:1867 Created:1868 Created:1869 Created:1870 Created:1871 Created:1872 Created:1873 Created:1874 Created:1875 Data

Genre:Correspondence Genre:Military documents

Person/Org:1st South Carolina Volunteer Infantry Regiment Person/Org:Adams, Charles Francis  
 Person/Org:Ball, James D. Person/Org:Benham, L.M. Person/Org:Bennett, William True  
 Person/Org:Chamberlain, George B. Person/Org:Child, Lydia Maria Francis  
 Person/Org:Devens, Charles Person/Org:Foster, Charles W. Person/Org:Hartwell, Alfred S.  
 Person/Org:Higginson, Thomas Wentworth Person/Org:Hinton, Ricard Josiah  
 Person/Org:Hyde, Edward W. Person/Org:Simpson, Wallace W. Person/Org:Saxton, Rufus  
 Person/Org:Stickles, Daniel Edgar Person/Org:Thompson, John Milton  
 Person/Org:Tremain, Henry Edwin Physical Format:Personal papers State:NA  
 Theme:African American Troops Theme:Black voices Theme:Union Theme:Women's voices  
 UID:hou00124c01193

 **hou00124c01193\_0001.jpg**  
 JPEG Image - 1.3 MB  
 Published Feb 17, 2023  
 10 Downloads  
 MDS: ac1..bcb

File associated with: Higginson, Thomas Wentworth, 1823-1911. Post-war correspondence, 1866-1875.  
 Origin of source: <https://nrs.harvard.edu/URN-3:FHCL.HOUGH:101293733>

City:Havana Country:United States Created:1866 Created:1867 Created:1868 Created:1869 Created:1870 Created:1871 Created:1872 Created:1873 Created:1874 Created:1875 Data

Genre:Correspondence Genre:Military documents

Person/Org:1st South Carolina Volunteer Infantry Regiment Person/Org:Adams, Charles Francis  
 Person/Org:Ball, James D. Person/Org:Benham, L.M. Person/Org:Bennett, William True  
 Person/Org:Chamberlain, George B. Person/Org:Child, Lydia Maria Francis  
 Person/Org:Devens, Charles Person/Org:Foster, Charles W. Person/Org:Hartwell, Alfred S.  
 Person/Org:Hinton, Ricard Josiah

Harvard Dataserve > Houghton Library Dataserve Collection > Slavery, Abolition, Emancipation, and Freedom Collection > hou00124c01193 >

**hou00124c01193\_0001.jpg**  
 This file is part of "hou00124c01193".

[View LL](#)

File Citation  
 Houghton Library, 2023. "hou00124c01193". <https://doi.org/10.7910/DVN/RFT3ME>. Harvard Dataserve, V1. UNF:6:CSV656rH4yq3NmkZVbe= [file:UNF]

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

Access File -  
 Contact Owner Share


File Metrics  
 10 Downloads

Dataset Citation  
 Houghton Library, 2023. "hou00124c01193". <https://doi.org/10.7910/DVN/RFT3ME>. Harvard Dataserve, V1. UNF:6:CSV656rH4yq3NmkZVbe= [file:UNF]

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

Preview Metadata Versions

Open in New Window



Charles...  
 one B-85  
 Dear Sir  
 I have your note as to the Col. Grant's... He was arrested and tried by order of Maj. Gen. Gileman - You are mistaken in supposing that the man executed by him was either a guerrilla or assassin - He had been a rebel soldier and was returning to his home in Va.

# 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

Dataset Metrics

15,381 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.

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 15 Files Download Request Access
- GSP\_DataUse\_Terms\_140422.pdf
  - Adobe PDF - 282.7 KB
  - Published Aug 24, 2014
  - 1,649 Downloads
  - MDS: e64...059
  - Data use terms
- GSP\_list\_140630.csv
  - Plain Text - 636.4 KB
  - Published Aug 24, 2014
  - 915 Downloads
  - MDS: 27e...4a8
  - Demographic, cognitive/behavior, quality control, and morphometrics data.
- GSP\_part10\_140630.tar
  - TAR Archive - 10.0 GB
  - Published May 21, 2014
  - 1,239 Downloads
  - MDS: 302...48e
  - Tar archive of imaging data for subjects 1414-1570; refer to GSP\_README\_140630.pdf for more information.
- GSP\_part1\_140630.tar
  - TAR Archive - 10.1 GB
  - Published May 21, 2014
  - 1,284 Downloads
  - MDS: a33...f55

### Citation Metadata

**Persistent Identifier** doi:10.7910/DVN/25833

**Publication Date** 2014-05-22

**Title** Brain Genomics Superstruct Project (GSP)

**Author** Buckner, Randy L. (Harvard University)  
Roffman, Joshua L.  
Smoller, Jordan W.

**Point of Contact** Use email button above to contact.

[GSP Data Release \(Harvard University\)](#)

### Description

### Related Material

Examples of prior publications of GSP data with partial data description:

Yeo, B.T., Krienen, F.M., Sepulcre, J., Sabuncu, M.R., Lashkari, D., Hollinshead, M., Roffman, J.L., Smoller, J.W., Zolnick, J.R., Fischl, B., Liu, H., Buckner, R.L. (2011) The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology*, 106(3): 1125-1165. [Link to article](#)

Buckner, R.L., Krienen, F.M., Castellanos, A., Diaz, J.C., Ye o, B.T. (2011) The organization of the human cerebellum and intrinsic functional connectivity. *Journal of Neurophysiology*, 106(5): 2322-2345. [Link to article](#)

**Access Dataset**

Contact Owner Share

**Dataset Metrics**

15,385 Downloads

### Dataset Terms

#### License/Data Use Agreement

Our [Community Norms](#) as well as good scientific practices expect that proper credit is given via citation. Please use the data citation shown on the dataset page.

Custom Dataset Terms – the following Custom Dataset Terms have been defined for this dataset.

#### Terms of Use

**These restrictive terms of use take precedence over any less restrictive use terms that apply generally to Dataverse Network Terms of Use**

- I request access to data collected as part of the Brain Genomics Superstruct Project (GSP) of Harvard University and the Massachusetts General Hospital, and I agree to the following:
  - I will not attempt to establish the identity of or attempt to contact any of the included human subjects.
  - I will not attempt to link any of the distributed data to any other data that might contain information about the included human subjects.
  - I understand that under no circumstances will the code that would link these data to Protected Health Information be given to me, nor will any additional information about individual human subjects be released to me under these Open Access Data Use Terms.
  - I will comply with all relevant rules and regulations imposed by my institution. This may mean that I need my research to be approved or declared exempt by a committee that oversees research on human subjects e.g., my Inter nal Review Board or Ethics Committee. Different committees operate under different national, state, and local laws and may interpret regulations differently, so it is important to ask about this.
  - I may redistribute original GSP Open Access data and any derived data as long as the data are redistributed under these same Data Use Terms.
  - I will acknowledge the use of GSP data and data derived from GSP data when publicly presenting any results or algorithms that benefitted from their use.
  - Papers, book chapters, books, posters, oral presentations, and all other printed and digital presentations of results derived from GSP data should contain the following word and/or in the acknowledgments section: *“Data were provided [in part] by the Brain Genomics Superstruct Project of Harvard University and the Massachusetts General Hospital. (Principal Investigators: Randy Buckner, Joshua Roffman, and Jordan Smoller), with support from the Center for Brain Science Neuroinformatics Research Group, the Athinoua A. Martineau Center for Biomedical Imaging, and the Center for Human Genetic Research. 20 individual investigators at Harvard and MGH generously contributed data to GSP Open Access Data Use Terms Version: 2014-Apr-22 the overall project.”*
  - Authors of publications or presentations using GSP data should cite relevant publications describing the methods used by the GSP to acquire and process the data. The specific publications that are appropriate to cite in any given study will depend on what GSP data were used and for what purposes. An annotated and appropriately up-to-date list of publications that may warrant consideration is available at <http://neuroinformatics.harvard.edu/gsp/>
  - The GSP as a consortium should not be included as an author of publications or presentations if this authorship would be based solely on the use of GSP data.
  - Failure to abide by these guidelines will result in termination of my privileges to access GSP data.

#### Restricted Files + Terms of Access

#### Restricted Files

There are 13 restricted files in this dataset.

#### Request Access

Users may request access to files.

# Replication Data for: Designing Social Inquiry: Scientific Inference for Qualitative Research

Version 1.0



Keohane, Robert O.; King, Gary, 2021, "Replication Data for: Designing Social Inquiry: Scientific Inference for Qualitative Research", <https://doi.org/10.7910/DVN/YHZG5M>, Harvard Dataverse, V1, UNF:6:2HECL90TQxdW2/NYMrkg== [fileUNF]

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

Access Dataset -

Contact Owner Share

Dataset Metrics

69 Downloads

## Description

Replication data for the preface of Gary King, Robert O. Keohane, and Sidney Verba. Designing Social Inquiry: Scientific Inference in Qualitative Research, 2nd Edition. Princeton University Press. Princeton, 2021. (2021-01-05)

## Subject

Social Sciences

## Keyword

Meta-analysis, qualitative research, literature review

## Related Publication

Gary King, Robert O. Keohane, and Sidney Verba. Designing Social Inquiry: Scientific Inference in Qualitative Research, 2nd Edition. Princeton University Press. Princeton, 2021.

## License/Data Use Agreement



Files Metadata Terms Versions

1 File



KKV2Sample (1).tab

Tabular Data - 279 B  
Published Jan 5, 2021  
69 Downloads

15 Variables, 5 Observations UNF:6:2HECL...kbg==

Access File

File Access



Public

Download Options

Comma Separated Values (Original File Format)

Tab-Delimited

RData

Download Metadata

Variable Metadata

Data File Citation

Explore Options

Data Explorer v2

Export Metadata -

OAI\_ORE

DataCite

OpenAIRE

Schema.org JSON-LD

DDI

Dublin Core

Croissant

DDI HTML Codebook

JSON

Ask the Data



File Tools - Open in New Window

Ask the Data

View Data

What do you want to know?

What do you want to know?

Answer please

This data appears to be election data showing the results of a political race. The columns seem to represent different candidates running for a position, and the rows contain information about the votes each candidate received in different regions. The data includes the candidate names, their party affiliations, the number of votes they received in various regions, and the percentage of votes they obtained. Additionally, there is a column indicating the number of votes not cast in the race.

	Gray Davis*	Bill Simon	Reinhold Gulke	Peter Miguel Camejo	Gary David Copeland	Iris Adam	Anselmo A. Chavez	Will B. King	Rob Marink
	DEM	REP	AI	GRN	LIB	NL	(W/I)	(W/I)	(W/I)
Alameda	216,058	76,407	3,622	37,919	6,558	3,319	2	0	1
Percent	62.9%	22.3%	1.0%	11.0%	1.9%	0.9%	0.0%	0.0%	0.0%
Alpine	229	247	17	40	15	12	0	0	0
Percent	40.9%	44.2%	3.0%	7.2%	2.6%	2.1%	0.0%	0.0%	0.0%
Amador	4,437	6,997	338	740	246	220	0	0	0
Percent	34.2%	54.0%	2.6%	5.8%	1.8%	1.6%	0.0%	0.0%	0.0%
Butte	19,437	32,706	1,497	5,963	1,050	802	0	0	0
Percent	31.6%	53.3%	2.4%	9.7%	1.7%	1.3%	0.0%	0.0%	0.0%

Viewing rows 1 through 13 of 182

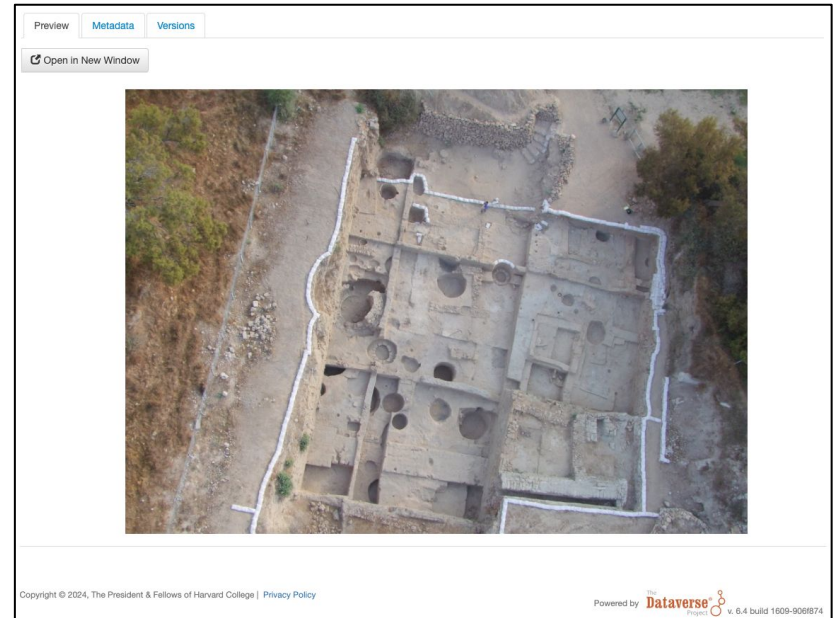
# Dataverse Tools and Features

# Dataverse Community Integrations & Tools

- **Tools that talk to Dataverse**
  - generally used to deposit data into Dataverse (via Deposit API)
  - usually don't require anything special to be set up in the Dataverse repository
- **Tools that Dataverse talks to**
  - user starts on Dataverse and is directed to the external tool
    - require manifest files
  - have predefined areas in the UI where these would plug into (**Explore** tools)
  - **OR**, are embedded into the Dataverse UI directly (**Preview** tools and **Query** Tools)
- **Tools that do both**
  - user starts on Dataverse and is directed to the external tool
    - require manifest files
  - also have predefined areas in the UI where these would plug into (**Configure** tools)
  - will also send something back to Dataverse, so need an API token that has "write" privileges

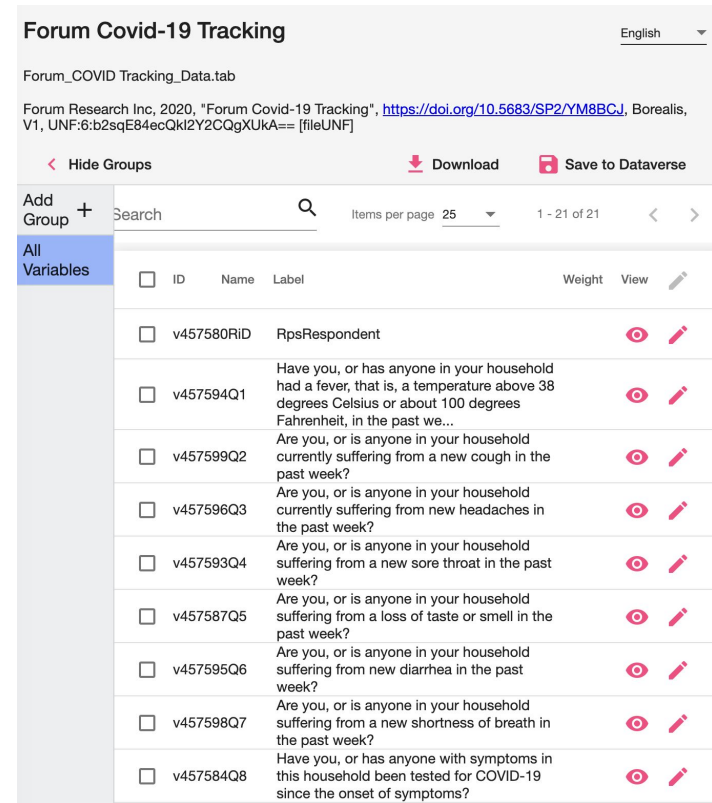
# File Previewers

- 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
  - GeoJSON
  - Zip files
  - NcML files
- Previewers are available through the preview (eye) icon on Dataset pages
- And also embedded as a tab on Datafile pages



# File Exploration, Configuration, and Query Tools

- File level **explore** tools provide a variety of features from data visualization to statistical analysis
- 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
- File level **configure** tools allow (authorized) users to send metadata about the file back to Dataverse

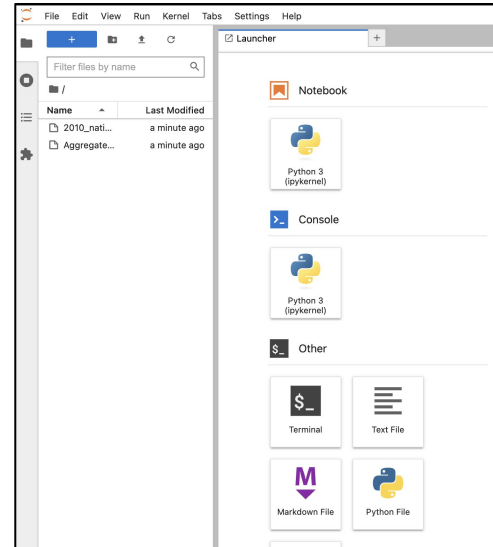
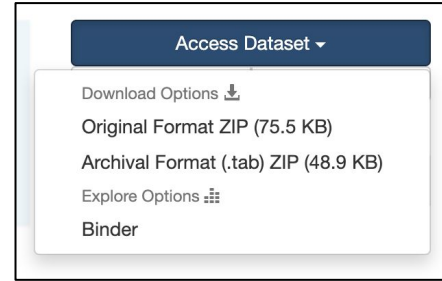


The screenshot displays the 'Forum Covid-19 Tracking' interface. At the top, it shows the title 'Forum Covid-19 Tracking' and the language 'English'. Below this, the file name 'Forum\_COVID Tracking\_Data.tab' is listed, along with a copyright notice for Forum Research Inc. (2020) and a DOI link. Action buttons for 'Hide Groups', 'Download', and 'Save to Dataverse' are visible. A search bar and 'Items per page' dropdown (set to 25) are also present. The main content is a table of variables, with the 'All Variables' tab selected. Each row includes a checkbox, an ID, a name, a label, a weight, and a view icon.

<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** - uses Open AI's ChatGPT & ICPSR best practices to provides recommendation to enhance metadata & generate meaningful titles, descriptions, and keywords





# Croissant

 Export Metadata ▾

OAI\_ORE

DataCite

OpenAIRE

Schema.org JSON-LD

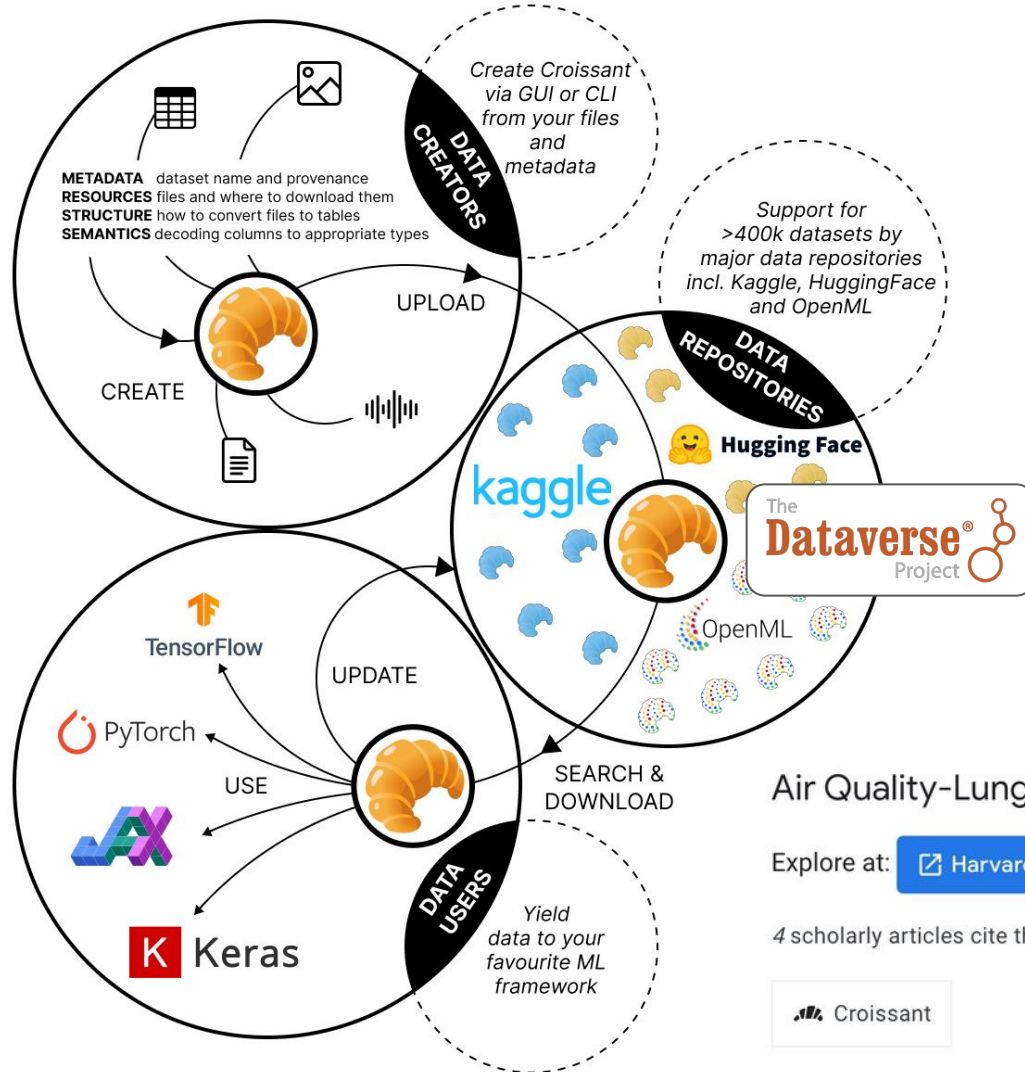
DDI

Dublin Core

**Croissant**

DDI HTML Codebook

JSON



Published September 26, 2024 | Version v1

Presentation

Open

# Building natural language interface for Dataverse network based on Croissant ML standard

Vyacheslav Tykhonov (Researcher)<sup>1</sup> 

Show affiliations

The new distributed network vision for AI is to create a shared data interface by querying multiple data nodes simultaneously. This distributed approach is based on the agile, community-driven standard called Croissant ML, allowing users to query and understand responses from various data platforms that support the standard, such as Dataverse, Kaggle, HuggingFace, and OpenML. The results are processed in a standardized way, converting metadata into a knowledge graph integrated with ontologies, and ingesting structured content into Large Language Models (LLMs) that act as reasoning engines, interfacing between humans and AI. The prototype enables "chatting" with individual data nodes in the network while considering ethical and privacy constraints, using only information shared through open metadata records.

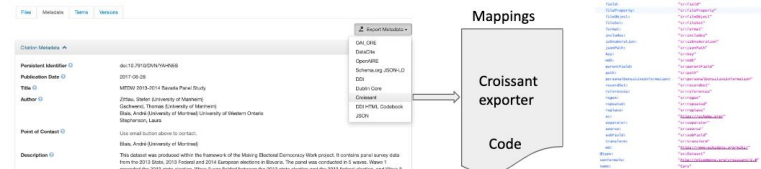
## Files

<https://zenodo.org/records/13842869>

Test it out:

<https://ai.muse-it.eu/?url=https://dataverse.jpl.nasa.gov/dataset.xhtml?persistentId=doi:10.48577/jpl.HEMWZF>

## Croissant ML export in Dataverse



## Experimental (graph) Croissant transformations in pyDataverse

```

4 Dataverse to Croissant.pyrb
File Edit View Insert Runtime Tools Help Croissant-dataverse
[+] Code Test Copy to Drive
[4] #!pip install -q pydataverse
#!pip install --upgrade --no-cache-dir git+https://github.com/zenodo-org/dataverse-croissant@pydataverse
import sys
host = "https://dataverse.jpl.nasa.gov"
pip = "pip install croissant"
croissant = Croissant(host, PDB)

#!pip3 install netlib
  
```

Local intelligence built on Dataverse + Llama3:8b + Graph

## Dataverse language Interface

Enter your question or DOI of some dataset, for example, doi:10.7910/DVN/XLJGUI

Question: Can you find a 2013 study on environmental policies in Bavaria?

Get Response

Click "Chat" button if you want to chat with some dataset

type:"dataset" AND (description:"study" OR title:"environmental policies" OR keyword:"environmental policies" OR description:"environmental policies" OR keyword:"study" OR title:"study") AND (description:"Bavaria" OR keyword:"Germany" OR keyword:"Bavaria" OR title:"Germany" OR title:"Bavaria" OR description:"Germany") AND (keyword:"2013" OR description:"2013" OR title:"2013")

[Chat] Zittlau, Stefan; Gschwend, Thomas; Blais, André; Stephenson, Laura, 2017, "MEDW 2013-2014 Bavaria Panel Study", <https://doi.org/10.7910/DVN/YAHN5S>, Harvard Dataverse, V1

Dataset [that contains the Croissant exports for each dataset in Harvard Dataverse](#)

**Metrics** 0 Downloads

The root dataverse.

Search this dataverse... **Advanced Search**

**Dataverses (35)**  
 **Datasets (15)**  
 **Files (0)**

**Dataset Type**  
Software (2)  
Dataset (1)

**Publication Year**  
2024 (3)

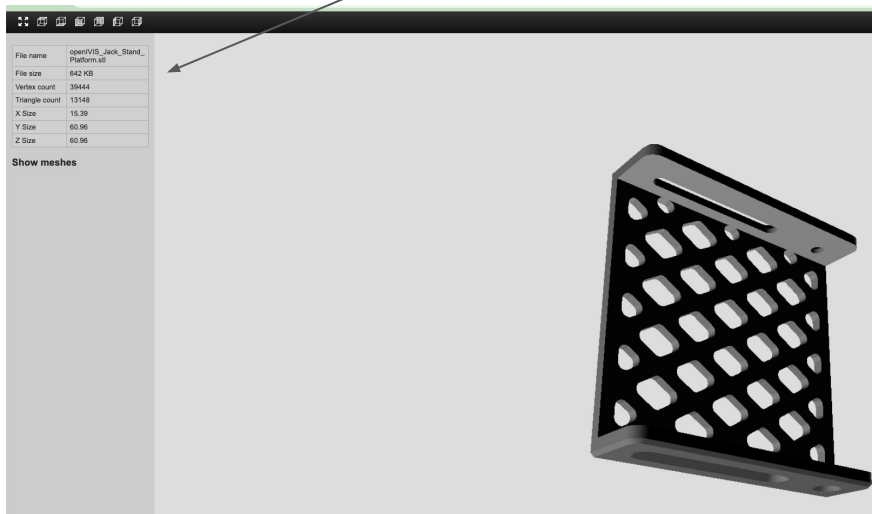
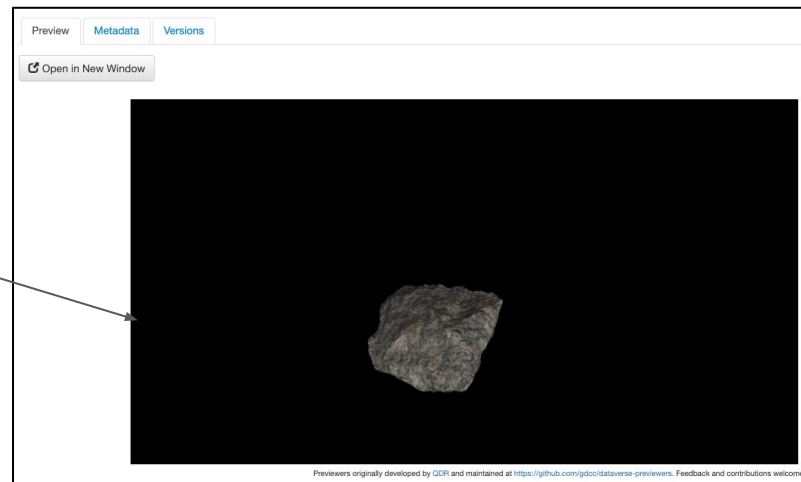
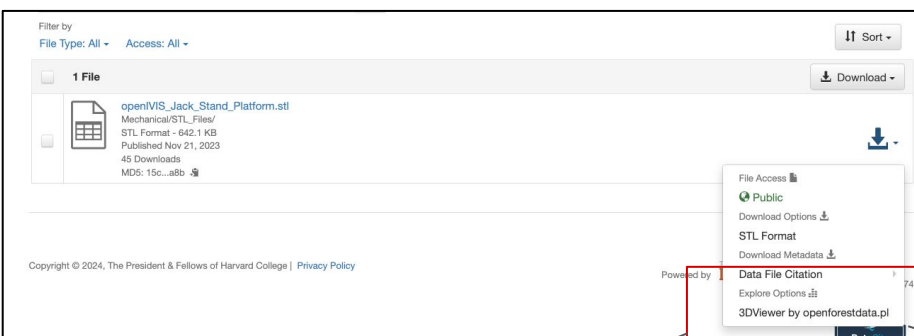
**Publication Status**  
Unpublished (47)

**1 to 10 of 50 Results**

**Awesome Dataset**  
Jul 31, 2024  
Admin, Dataverse, 2024, "Awesome Dataset", <https://doi.org/10.5072/FK2/LX1QBT>, Root, V1  
Dataset are still cool.

**pyDataverse** **Draft** **Unpublished**  
Jul 30, 2024 - dvf5499ba4  
Range, Jan, 2024, "pyDataverse", <https://doi.org/10.5072/FK2/30a7a5c>, Root, DRAFT VERSI  
A Python module for Dataverse.

<https://github.com/IQSS/dataverse/pull/10694>



The file in the screenshot:


<https://dataverse.harvard.edu/file.xhtml?fileId=7559161&version=1.1>

# Deluge of Large Data

2004 - 2023: < 70TB

Gen - Apr '24: + 35 TB (+50%)

Prospect by June '24: +420 TB

**Sustainability model:**   
data owner contributes to storage costs

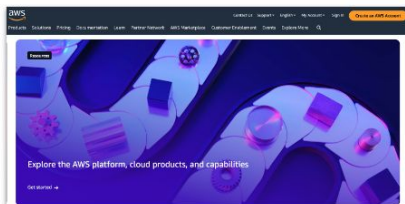
Moving to **MOC** (Mass Open Cloud):

- no egress costs
- 30% (storage/computing) cost of AWS



# Large data support infrastructure

Amazon Web Services (AWS) | 01/17/2017



New England  
Research Cloud

Northeast Storage Exchange (NESE)

Category	Storage Options	Description
Consultations	-	Basic and extended consultations to choose best service offering
Data Curation	-	Curation consultations and other services
Basic Service	NESE Tape NESE Disk	Dynamic file access for some files, delayed access for others
On-Demand Service	NESE Disk w/S3 AWS S3	Dynamic file access option using AWS S3 or NESE disk w/S3 mounted containers
Cold Storage	NESE Tape AWS Glacier	Cold storage for rarely accessed data
Custom	Combination	Custom solutions using multiple storage and access options



# Computing on the Data

# Computing on Data in Harvard Dataverse

**Need:** Researchers want to compute on HDV datasets

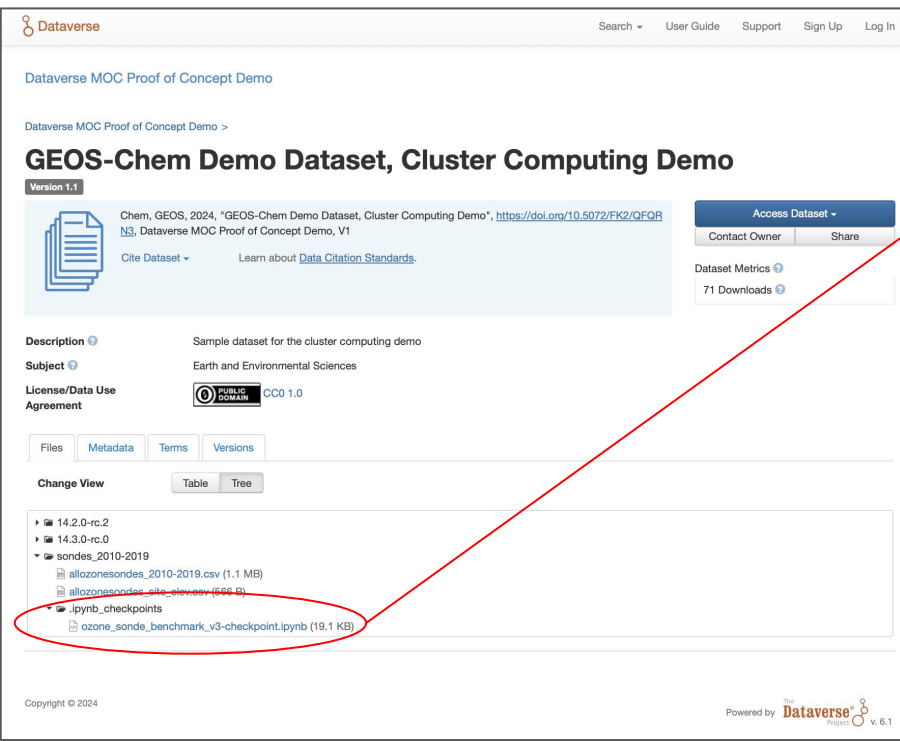
## Challenges

May not be possible to:

- Move/copy data files from HDV data storage
- Perform difficult computations using local resources

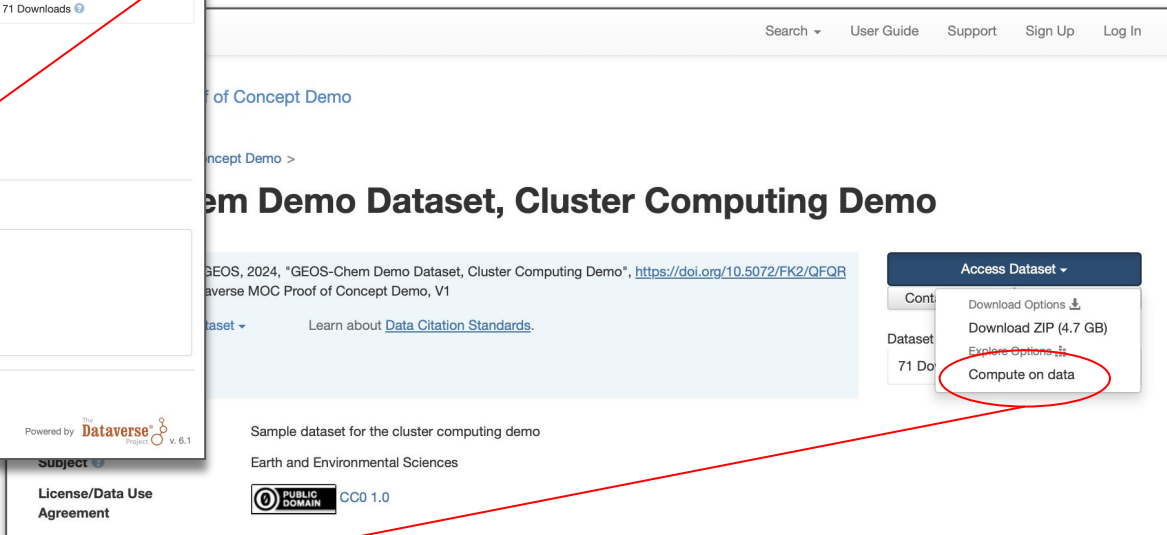
## Approaches

1. Run containerized Dataverse & compute where the data is stored on HPC (current PoC)
2. DV - Open OnDemand Integration: Move copy of files to HPC cluster (PoC in June 2025)



# PoC: Run containerized Dataverse on HPC

If Dataverse sees a (python) **notebook**, the new menu item **“Compute on data”** appears in the *Access Dataset* drop down menu



This **menu will launch** the JupyterLab VM with the pre-loaded notebook taken from the dataset. All files in this collection are seen as **local** to the Jupyter instance. Python will simply load them into memory for computing purposes.

# Computing on the data

jupyter dataverse\_geos\_chem Last Checkpoint: yesterday

File Edit View Run Kernel Settings Help Trusted

Markdown

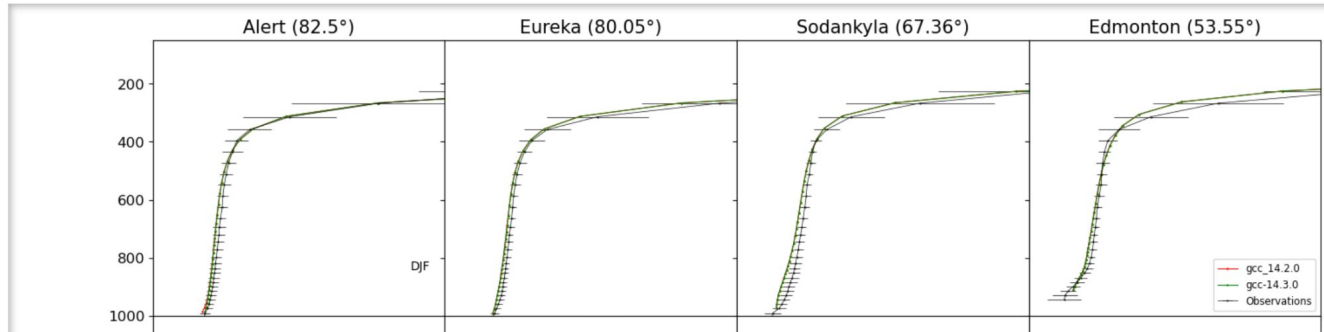
```
#pdf_pages.close()
```

## Create plots of GEOS-Chem model output vs. observational data

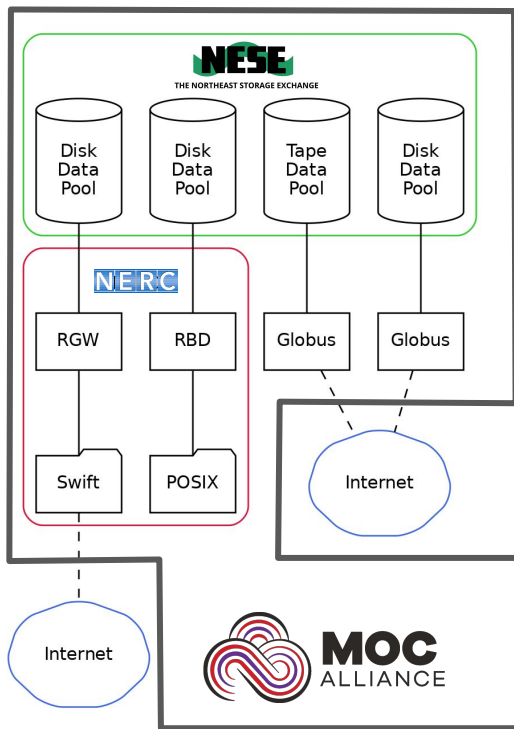
A new figure will be generated showing GEOS-Chem data vs. ozonesonde sites for each meteorological season (winter = DJF, spring = MAM, summer = JJA, fall = SON). Close each plot window to move to the next plot.

```
[6]: # Finally, generate plots of GEOS-Chem data vs. ozonesondes
# These are used for GEOS-Chem benchmarking and model evaluation
make_benchmark_models_vs_sondes_plots(
    obs_data_file,
    obs_site_file,
    ref_filepaths,
    ref_label,
    dev_filepaths,
    dev_label,
)
```

Then some nice computation happens



# Connecting Dataverse to HPC infrastructure:



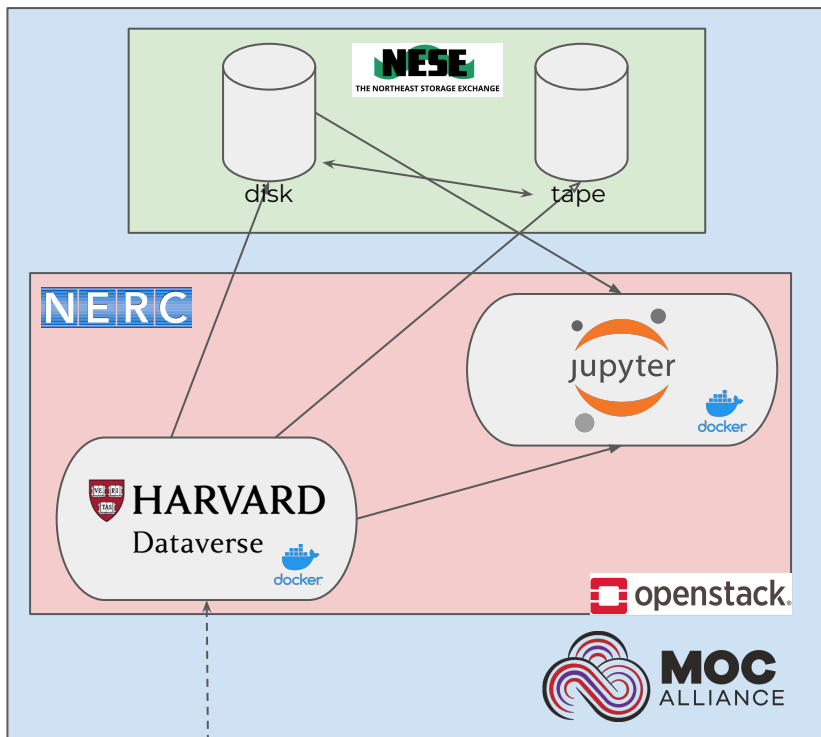
NESE — NERC Data Storage Stack



The screenshot shows the Dataverse web interface for a collection titled "Dataverse MOC Proof of Concept Demo".

- Header:** Dataverse logo and a hamburger menu icon.
- Collection Title:** Dataverse MOC Proof of Concept Demo
- Metrics:** A bar chart icon, "Metrics", and "71 Downloads".
- Actions:** "Contact" and "Share" buttons.
- Description:** "This collection is for the proof-of-concept demo of the MOC-Dataverse collaboration on the NERC cluster".
- Search:** A search bar with the text "Search this dataverse..." and a magnifying glass icon, followed by a link to "Advanced Search".
- Filters:** A list of filters: "Dataverses (0)", "Datasets (1)", and "Files (27)".
- Results:** A section titled "1 to 1 of 1 Result" with a "Sort" dropdown menu. The result is a dataset titled "GEOS-Chem Demo Dataset, Cluster Computing Demo" with a document icon and a date of "Feb 12, 2024". The description includes: "Chem, GEOS, 2024, "GEOS-Chem Demo Dataset, Cluster Computing Demo", <https://doi.org/10.5072/FK2/QFQRN3>, Dataverse MOC Proof of Concept Demo, V1". Below the description is the text "Sample dataset for the cluster computing demo".
- Metadata:** Fields for "Publication Year" (2024 (1)), "Author Name" (Chem, GEOS (1)), "Subject" (Earth and Environmental Sciences (1)), and "Deposit Date" (2024 (1)).

# Connecting Dataverse to HPC infrastructure: MOC



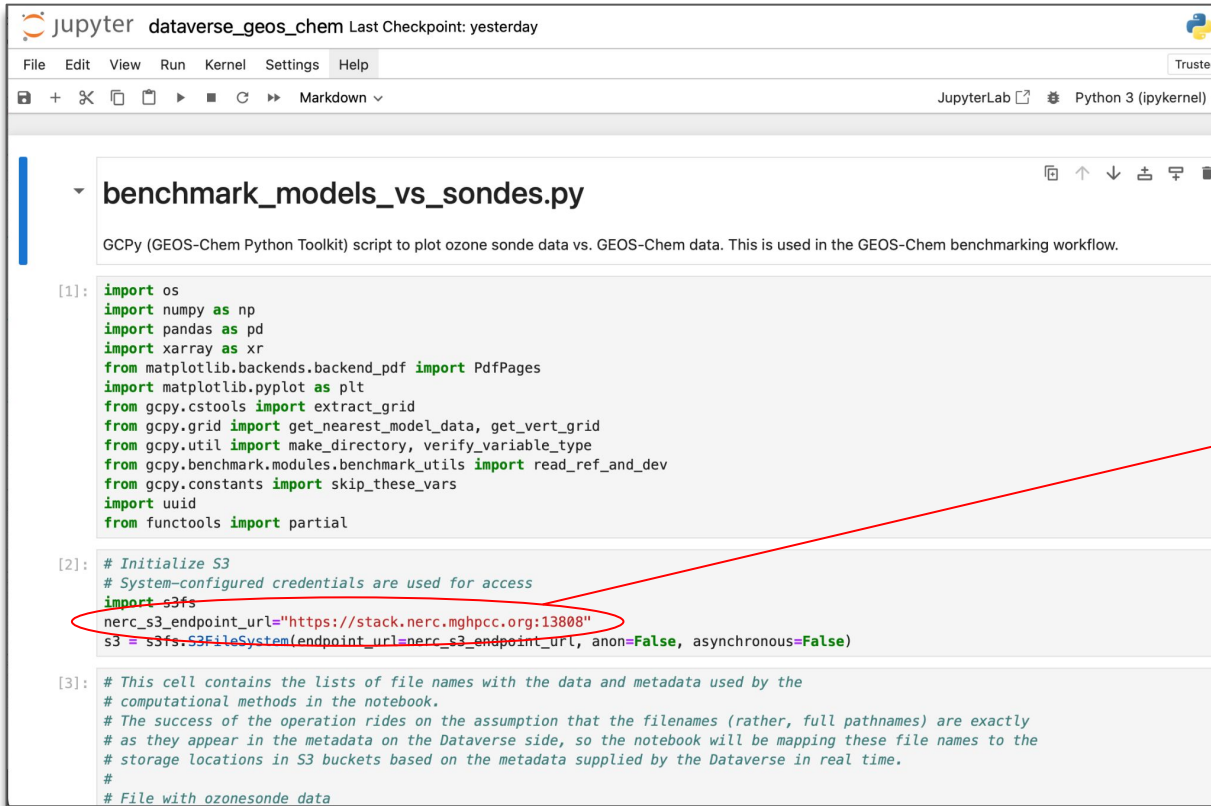
happy Harvard Dataverse user  
(only NERC PI's will be able to run compute)

The screenshot shows the Dataverse web interface. The page title is "Dataverse MOC Proof of Concept Demo". It displays metrics for 72 Downloads and a "Contact" button. A search bar is present with the text "Search this dataverse...". Below the search bar, there are filters for "Dataverses (0)", "Datasets (2)", and "Files (28)". The search results show two items:

- GEOS-Chem Dataset, Cluster Computing Demo**  
Feb 23, 2024  
Iacus, Stefano, 2024, "GEOS-Chem Dataset, Cluster Computing Demo", <https://doi.org/10.7910/MOC/7XKHQD>, Dataverse MOC Proof of Concept Demo, V1, UNF:6:mSD5ia+q1xk001hnBKJ5wQ== [fileUNF]  
Sample dataset for the cluster computing demo.
- GEOS-Chem Dataset, NESE Tape Storage Demo**  
Feb 23, 2024  
Admin, Dataverse, 2024, "GEOS-Chem Dataset, NESE Tape Storage Demo", <https://doi.org/10.7910/MOC/TOXECN>, Dataverse MOC Proof of Concept Demo, V1  
This dataset demonstrates the concept of Dataverse using managed remote storage volume, in this case with the data files stored on tape at NESE, accessible via Globus.

Copyright © 2024  
Powered by **Dataverse** Project  
v. 6.1

# Computing on the data



```

Jupyter dataverse_geos_chem Last Checkpoint: yesterday
File Edit View Run Kernel Settings Help Trusted
+ ✂ 📄 📄 ▶ ■ ↻ ▶▶ Markdown v JupyterLab Python 3 (ipykernel)

benchmark_models_vs_sondes.py
GCPy (GEOS-Chem Python Toolkit) script to plot ozone sonde data vs. GEOS-Chem data. This is used in the GEOS-Chem benchmarking workflow.

[1]: import os
import numpy as np
import pandas as pd
import xarray as xr
from matplotlib.backends.backend_pdf import PdfPages
import matplotlib.pyplot as plt
from gcpy.cstools import extract_grid
from gcpy.grid import get_nearest_model_data, get_vert_grid
from gcpy.util import make_directory, verify_variable_type
from gcpy.benchmark.modules.benchmark_utils import read_ref_and_dev
from gcpy.constants import skip_these_vars
import uuid
from functools import partial

[2]: # Initialize S3
# System-configured credentials are used for access
import s3fs
nerc_s3_endpoint_url="https://stack.nerc.mghpcc.org:13808"
s3 = s3fs.S3FileSystem(endpoint_url=nerc_s3_endpoint_url, anon=False, asynchronous=False)

[3]: # This cell contains the lists of file names with the data and metadata used by the
# computational methods in the notebook.
# The success of the operation rides on the assumption that the filenames (rather, full pathnames) are exactly
# as they appear in the metadata on the Dataverse side, so the notebook will be mapping these file names to the
# storage locations in S3 buckets based on the metadata supplied by the Dataverse in real time.
#
# File with ozonesonde data

```

NERC S3 endpoint for the containerized storage (which exists on NESE)

# Computing on the data

Automatic mapping of local file names (local to the python notebook) to Harvard Dataverse file pointers on NESE

```
mapping file pathnames as listed in the dataset metadata to the direct storage locations:  
sondes_2010-2019/allozonesondes_2010-2019.csv -> s3://mopcpcdemocontainer/10.5072/FK2/QFQRN3/18d8fa95075-22edad4355c8  
sondes_2010-2019/allozonesondes_site_elev.csv -> s3://mopcpcdemocontainer/10.5072/FK2/QFQRN3/18d8fa950f3-ed27dd9b2f0a  
14.3.0-rc.0/GCCClassic/FullChem/OutputDir/GEOSChem.SpeciesConc.20190101_0000z.nc4 -> s3://mopcpcdemocontainer/10.5072/FK2/QFQRN3/18d8fa81861-a21e0a10e8a5  
14.2.0-rc.2/GCCClassic/FullChem/OutputDir/GEOSChem.SpeciesConc.20190101_0000z.nc4 -> s3://mopcpcdemocontainer/10.5072/FK2/QFQRN3/18d8fa4ee6e-12fc2fa05bf8  
14.2.0-rc.2/GCCClassic/FullChem/OutputDir/GEOSChem.SpeciesConc.20190201_0000z.nc4 -> s3://mopcpcdemocontainer/10.5072/FK2/QFQRN3/18d8fa4cd02-4a684a55b7dc  
14.3.0-rc.0/GCCClassic/FullChem/OutputDir/GEOSChem.SpeciesConc.20190201_0000z.nc4 -> s3://mopcpcdemocontainer/10.5072/FK2/QFQRN3/18d8fa7ee17-5242cdf4a74  
14.2.0-rc.2/GCCClassic/FullChem/OutputDir/GEOSChem.SpeciesConc.20190301_0000z.nc4 -> s3://mopcpcdemocontainer/10.5072/FK2/QFQRN3/18d8fa53f9b-7d6cae9694b7  
14.3.0-rc.0/GCCClassic/FullChem/OutputDir/GEOSChem.SpeciesConc.20190301_0000z.nc4 -> s3://mopcpcdemocontainer/10.5072/FK2/QFQRN3/18d8fa85f90-048bd360915f  
14.2.0-rc.2/GCCClassic/FullChem/OutputDir/GEOSChem.SpeciesConc.20190401_0000z.nc4 -> s3://mopcpcdemocontainer/10.5072/FK2/QFQRN3/18d8fa4f6fe-b811828db8aa  
14.3.0-rc.0/GCCClassic/FullChem/OutputDir/GEOSChem.SpeciesConc.20190401_0000z.nc4 -> s3://mopcpcdemocontainer/10.5072/FK2/QFQRN3/18d8,
```

An upcoming version of **pyDataverse** will automatically generate these mappings for general cloud systems.

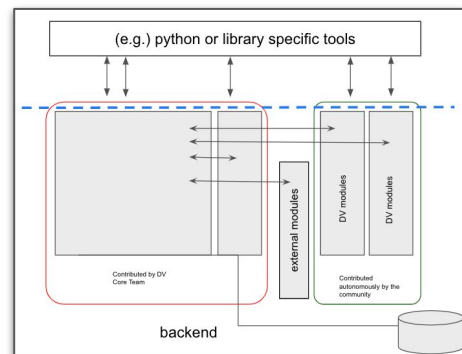
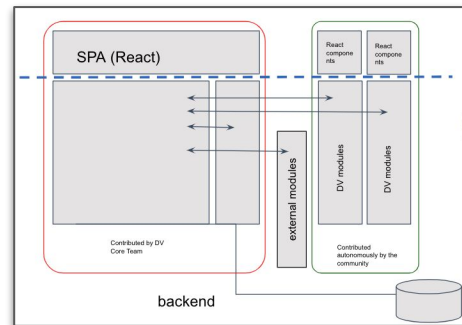
# The re-arch project (last 12 months)

## Goals

- **Modernize** the application
- **Separate backend & frontend**
- Increase **interoperability**
- **Modularize** backend & frontend
- **Speed up development**
- Deploy **new UI/UX ideas**
- **Shorten release time**
- **Empower** the **community**

## Approach

- Single Page Application (**SPA**)
- Native support for **localization**
- Native support for **accessibility**
- **Full automatic testing**
- Improved **modularity**
- **“Plug&play”** module approach
- **Lighter** modules
- Optimization of backend



# Grants and Collaborations

# NIH Generalist Repository Ecosystem Initiative (GREI)

The mission of GREI is to establish a common set of capabilities, services, metrics, and social infrastructure; raise general awareness and facilitate researchers to adopt FAIR principles to better share and reuse data.

This initiative will further enhance the biomedical data ecosystem and help researchers find and share data from NIH-funded studies in generalist repositories.

## Goals of the Generalist Repository Ecosystem Initiative



1

Make it easier for researchers to **share data**.



2

Enable the improved **discoverability** of NIH-funded data across generalist repositories.



3

Support greater **reproducibility** of NIH-funded research by ensuring data associated with publications is readily available.



4

**Avoid duplication** of the data across repositories.



5

Encourage NIH-funded researchers to be both contributors and consumers to **increase the reuse** of data.

# NIH GREI (2022-ongoing)

## NIH GREI (Generalist Repository Ecosystem Initiative)

- Current programs and activities involved:
  - Remote Large Storage Support
  - Controlled Vocabularies for Biomedical
  - Discovery for Metadata (e.g. Data Documentation Initiative (DDI))
  - Computational Workflows
  - Harvesting and Sharing Metadata Across Repositories
  - Usage Metrics - Make Data Count Support
  - Improvements for Sensitive Data Support
  - Evaluation and Evolution of Architecture
  - NIH Data Management Plans
  - Community Repository Training
  - Support for Research Objects (beyond datasets)

## Participating Repositories





## Current Features



**Installations**



**Collections**



**Datasets**



**Files**

## NIH OTA Adds...



### Datasets

- New Workflows for **Large Data Support** ([Globus](#), [demo](#))
- **Remote** Data Support ([TRSA](#))
- Flexible **Biomedical Metadata** Support through External Vocabularies and Data Dictionaries (UMLS, MeSH)
- Investigate CEDAR integration



### Files

- Additional Metadata for **Code files** ([Codemeta](#))
- Support for **Replication Packages** through Workflows and Containers
- **Sensitive Data Support** through Differential Privacy ([OpenDP](#), [DataTags](#), [PSIprivacy](#))
- Encryption

- Usage Metrics
- UX/UI Enhancements
- New APIs
- Interoperability
- Metadata Harvesting
- Curation Services
- [Training and Outreach](#)

# ORCID Global Participation Grant awarded to GDCC

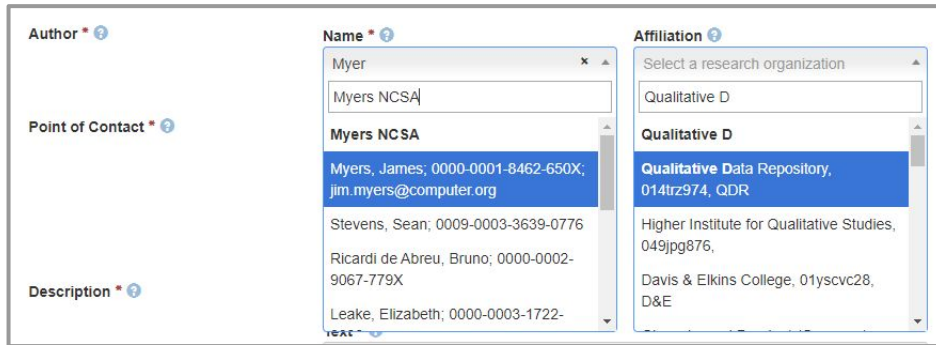
- Working with the Dataverse Community and Global South Dataverse Instances to:
  - Deploy current ORCID/ROR functionality
  - Verify/Promote addition of published datasets to ORCID profiles (via DataCite DataCommons)
  - Design/test new DOI lookup for Related Publications
  - Present results to the Dataverse Community

\*\* If you're interested in deploying the existing ORCID/ROR capabilities and/or in helping to design/test new ORCID-related functionality, please contact Jim Myers ([qqmyers@hotmail.com](mailto:qqmyers@hotmail.com))

# Dataverse 6.4 ORCID/ROR Support

- Lookup of People and Organizations via ORCID and ROR
- Can be configured for Author (name/affiliation), Depositor, Funding Information/Agency and other fields

- Edit:



The screenshot shows the 'Edit' form for an author. The 'Name' dropdown menu is open, showing a search for 'Myer' with results including 'Myers NCSA' (selected), 'Myers, James; 0000-0001-8462-650X; jim.myers@computer.org', 'Stevens, Sean; 0009-0003-3639-0776', 'Ricardi de Abreu, Bruno; 0000-0002-9067-779X', and 'Leake, Elizabeth; 0000-0003-1722-'. The 'Affiliation' dropdown menu is also open, showing a search for 'Qualitative D' with results including 'Qualitative D', 'Qualitative Data Repository, 014trz974, QDR' (selected), 'Higher Institute for Qualitative Studies, 049jpg876', and 'Davis & Elkins College, 01yscvc28, D&E'.

Start typing a name, acronym, or other information from a profile

Select from matching values

- Dataset Page:



The screenshot shows the 'Author' field on a dataset page. The text reads 'Myers, James  (Qualitative Data Repository )'.

Display shows name and a live link to the ORCID/ROR profile

Exported metadata contains the ORCID/ROR id for machine processing

Try it now at <https://demo.dataverse.org>

New and Upcoming Feature...



## New or Upcoming Features & Services

### New

**Search and Facet by License:** New ability to search and facet by dataset license (e.g., CC0)

**New Contributor Guide:** The UX Working Group released a new Dataverse [Contributor Guide](#).

**File-Level Retention Periods:** Dataverse now supports new file-level retention periods. See the [Retention Periods section](#) of the guide for details.

**Support for Multiple Persistent Identifier Providers:** Support for using multiple PID provider accounts (e.g., DOI, Handle, PermaLink). Great for consortia.

### Forthcoming

**Traditional Knowledge Labels:** Support for datasets using the TK Labels controlled vocabulary

**Dataverse Marketplace:** Centralized location for simple setup of configurable plug ins such as exporters, external tools, etc.

**External Search:** New External Tool type, “search”, for external API-enabled search engines to search Dataverse content in different ways

**Dataverse Metrics Hub:** A new service providing metrics about the Dataverse Project, Harvard Dataverse, and the Dataverse network via APIs

# The Global Dataverse Community

# Global Research Community...

**ADA AUSTRALIAN DATA ARCHIVE**  
(Australian National University)

Search User Guide Support Sign Up Log In

ADA Datasets > ANU Poll Datasets >

## ANU Poll 47 (August 2021): Mental health and schooling of children during COVID 19

Version 1.0

Biddle, Nicholas, 2021, "ANU Poll 47 (August 2021): Mental health and schooling of children during COVID-19", <https://doi.org/10.26193/133C0X>, ADA Datasets

Cite Dataset Learn about Data Citation Standards

**Description** This is the 47 waves of data collected in Australia. The survey was run on behalf of the Australian Government, covering issues, including mental health and wellbeing.

**Subject** Social Sciences

**Keyword** Public Health, Children, Stress

**License/Data Use Agreement** Custom Dataset Terms

**HARVARD Dataverse**

## COVID-19 Data Collection

Harvard Dataverse >

This is a general collection of COVID-19 data deposited in the Harvard Dataverse repository. The list in this collection is maintained (IQSS and Harvard Library). Researchers who deposit their related data into Harvard Dataverse will have their data linked to this collection. Please use the [contact link](#) if you have any questions about this collection.

Search this dataverse... Advanced Search

1 to 10 of 2,561 Results

- Healthcare policies by Brazilian States during the Covid-19 pandemic**  
May 23, 2024 · Harvard Dataverse  
Bertholini, Frederico; Fernandez, Michelle; Maia, Barbara, 2023, "Healthcare policies by Brazilian States during the Covid-19 pandemic", <https://doi.org/10.7910/DVN/HF0KGG>, Harvard Dataverse, V4  
This dataset provides a comprehensive overview of the healthcare policies implemented by state governments in Brazil during the progression of the Covid-19 pandemic. To achieve this, we analyzed the regulatory frameworks of healthcare policies developed by the 27 Brazilian states...
- Longitudinal COVID-19 Survey**  
May 22, 2024 · Longitudinal COVID-19 Survey  
Ripberger, Joseph; Jenkins-Smith, Hank; Gupta, Kulkaj; Silva, Carol; Robinson, Scott; Ross, Jennifer; Fox, Andrew, 2024, "Longitudinal COVID-19 Survey", <https://doi.org/10.7910/DVN/MRBSNG>, Harvard Dataverse, V1, UNF:8:LtQV00u59WejyRf9YQ== [fileUNF]  
This repository contains data and metadata (survey instruments) from a longitudinal COVID-19 Survey conducted March 2020 - March 2021 by the Institute for Public Policy Research and Analysis at the University of Oklahoma. The survey, which consists of 13 waves, collected weekly a...
- Longitudinal COVID-19 Survey (University of Oklahoma)**  
May 22, 2024 · Harvard Dataverse

**odesi** Statistics Canada Public Use Microdata Files (PUMFs) / Collection de fichiers de microdonnées à grande diffusion (FMGD)

Odesi Search and Explore | Rechercher et explorer

Borealis > Odesi > Statistics Canada Public Use Microdata Files (PUMFs) / Collection de fichiers de microdonnées à grande diffusion (FMGD) >

## Impacts of COVID-19 on Canadians, 2020: Perceptions of Safety

Version 1.0

Statistics Canada, 2023, "Impacts of COVID-19 on Canadians, 2020: Perceptions of Safety", <https://doi.org/10.5683/SP3/CQ3FLA>, Borealis, V1, UNF:6:bFh3r7KzLIn3ZZhRCvAg== [fileUNF]

Cite Dataset Learn about Data Citation Standards

**Description** This survey includes information from the community during the COVID-19 pandemic.

**Subject** Social Sciences

**Keyword** COVID-19, Crime and justice

**License/Data Use Agreement** Custom Dataset Terms

**e-ciencia Datos** Geolocation Metrics Help English Log In

Repositorio de Datos URJC | Biblioteca URJC

e-cienciaDatos > Repositorio de Datos URJC >

## Dental emergency care in Spain during the state of alarm due to COVID-19 pandemic

Version 1.0

Escribano, Nuria; Giráldez, Isabel; Laura Ceballos; Cerdán, Fátima; Infante, Raquel; Fuentes, Mª Victoria, 2020, "Dental emergency care in Spain during the state of alarm due to COVID-19 pandemic", <https://doi.org/10.21950/SSTT2Q>, e-cienciaDatos, V1

Cite Dataset Learn about Data Citation Standards

**Description** The state of alarm due to COVID-19 in Spain led to limit dental treatment exclusively to emergencies. The objective of the survey was to evaluate the amount and type of emergencies attended during this period, as well as to know how they were solved and what measures were adopted to carry out dental care in these exceptional circumstances. (2020-11-17)

**Subject** Medicine, Health and Life Sciences

**Keyword** Dental emergencies, COVID-19, state of alarm, survey

**License/Data Use Agreement** PUBLIC DOMAIN CC0 1.0

Access to full dataset metrics

Visualizations	Downloads
1522	104

0 Citations from Crossref

# The State of the Dataverse Community is strong!

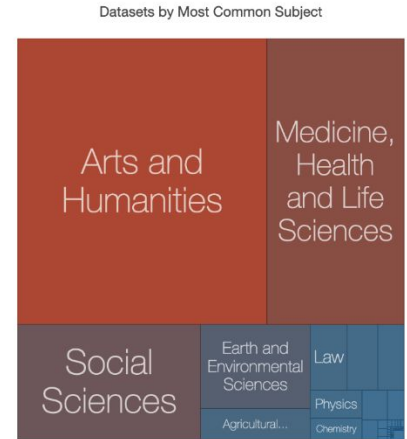
Dataverse Community (users, developers, researchers, librarians, data scientists and more)

- 120+ installations across the world
- 35 countries
- ~1000 subscribers to the Google Group
- ~100 people in chat.dataverse.org (Zulip)
- 253 “DV Community” Slack user
- 1600+ GitHub issues opened by the community
- 1300+ pull requests made by the community
- 185+ Github contributors to the main code base
- 186 videos on DataverseTV
- 189 community calls since 2016
- Many working groups, interest groups, community groups, meetings, etc.

The Data (dataverse.org/metrics)

- 123 installations
- 16,600 Dataverse Collections\*
- 426,000 Datasets\*
- 6,120,000 Files\*
- 86,500,000 File Downloads\*

\* metrics collected from 96 installations  
(running 4.9 and newer)

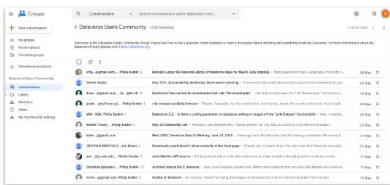


## Dataverse Google Group

### Recent Discussions:

- Technical setup & development questions
- Community calls & meetings
- Integrations & externals tools
- Preservation & FAIR Principles
- Community best practices
- Geospatial data integration
- User & Developer Guides
- Global Dataverse Community Consortium (GDCC)
- Surveys & community feedback

### Dataverse Google Group



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

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

## Annual [Dataverse Community Meeting](#)

- 2015 - IQSS, Harvard: [Common Models and APIs for Data Publishing and Citation Workshop](#)
- 2016 - Harvard Medical School, Harvard: [Fostering the Dataverse Community](#)
- 2017 - IQSS, Harvard: [Dataverse and the Data Lifecycle](#)
- 2018 - IQSS, Harvard: [Dataverse in the Age of Data Science](#)
- 2019 - IQSS, Harvard: Data Sharing: [Supporting Sustainability and Growth](#)

### 1st virtual meetings:

- 2020 - Global Virtual Conference: [Global COVID-19 Data Sharing](#)
- 2021 - Global Virtual Conference: [Automated CI/CD Testing, Installation and Deployment of a Dataverse Installation on a Cloud](#)
- 2022 - Global Virtual Conference: [Indigenous Data Sovereignty](#)

### 1st Regional Community Meetings:

- 2023 - University of Minho, Portugal: [Sharing data for future generations - sustainability, trust, and community building](#)
- 2024 - Texcoco, Mexico (CIMMYT): [Harvesting data from the field to the cloud](#)

# 10 years of community meetings!



[dataverse.org/events](https://dataverse.org/events)

# Dataverse Community Meeting 2025

June 9-12 at the University of  
North Carolina at Chapel Hill

**MORE INFORMATION COMING SOON!**

<https://dcm2025.hsites.harvard.edu/>

#Dataverse2025



# Thank You!



Sonia Barbosa  
*Associate Director, Dataverse, Data Curation,  
and The Murray Archive*