

Research software and

The
Dataverse[®]
Project 



HARVARD
Dataverse

Philip Durbin
[@philipdurbin](https://twitter.com/philipdurbin)

 **IQSS**
The Institute for Quantitative Social Science



What is Dataverse?



Open source research data repository software

Dataverse®

Dataverse is an [open source](#) web application for sharing, citing, analyzing, and preserving research data (developed by the [Data Science and Products 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](#).

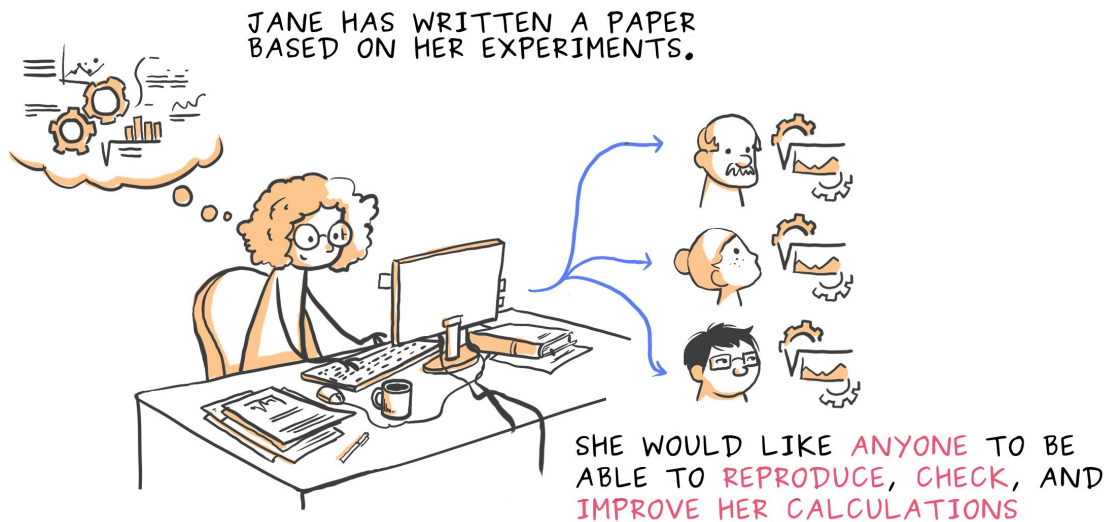
We love contributors! Please see our [Contributing Guide](#) for ways you can help.

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

Agenda



Code deposit



Reproducibility

Code already in Dataverse

Harvard Dataverse

5,700,913 Downloads

fileContentType:type/x-r-syntax

1 to 10 of 4,945 Results

- Dataverses (0)**
- Datasets (0)**
- Files (4,945)**

Metadata Source
Harvard Dataverse (4,872)
Harvested (73)

Publication Year
2018 (1,653)
2017 (1,514)
2016 (1,136)
2015 (382)

Simons Mallinson replication script.R
Jul 29, 2015 - Replication data for: Party Control and Perverse Effects in M...
Replication Challenges When Using DW-NOMINATE
R Syntax - 4.5 KB - MD5: ebd7238ab5983fcca8b5bb0926df949
[Code](#)
Script for reproducing analysis, tables, and figures from the published article

replication appendix.r
Mar 11, 2016 - Replication data for: Analyzing Manifestos in their Electora...
Applied to Austria, 2002&2008
R Syntax - 26.2 KB - MD5: 4b077329d83599e1f59cc779d9477a02

Harvard Dataverse

5,700,906 Downloads

Search this dataverse...

1 to 10 of 2,735 Results

- Dataverses (0)**
- Datasets (0)**
- Files (2,735)**

Metadata Source
Harvard Dataverse (2,701)
Harvested (34)

Publication Year
2018 (1,309)
2017 (548)
2016 (332)
2015 (292)
2014 (86)

File Type
Application (1,058)
Data (709)
Unknown (454)
Text (386)
Document (59)

File Tag
Code (2,735)
Documentation (65)
Data (61)
Simulation (21)
Coding Manual (20)

BA_estimation.R
Jan 18, 2019 - Replication Data for: When Should We Use Unit Fixed Effects Regression Models for Causal Inference with Longitudinal Data?
R Syntax - 14.9 KB - MD5: 54bab35e74ac0f65d4b38bbab7575c11
[Code](#)

DyadFE_estimation.R
Jan 18, 2019 - Replication Data for: When Should We Use Unit Fixed Effects Regression Models for Causal Inference with Longitudinal Data?
R Syntax - 10.3 KB - MD5: 59b86b8e8d3956dc6b4f6ea64110f396
[Code](#)

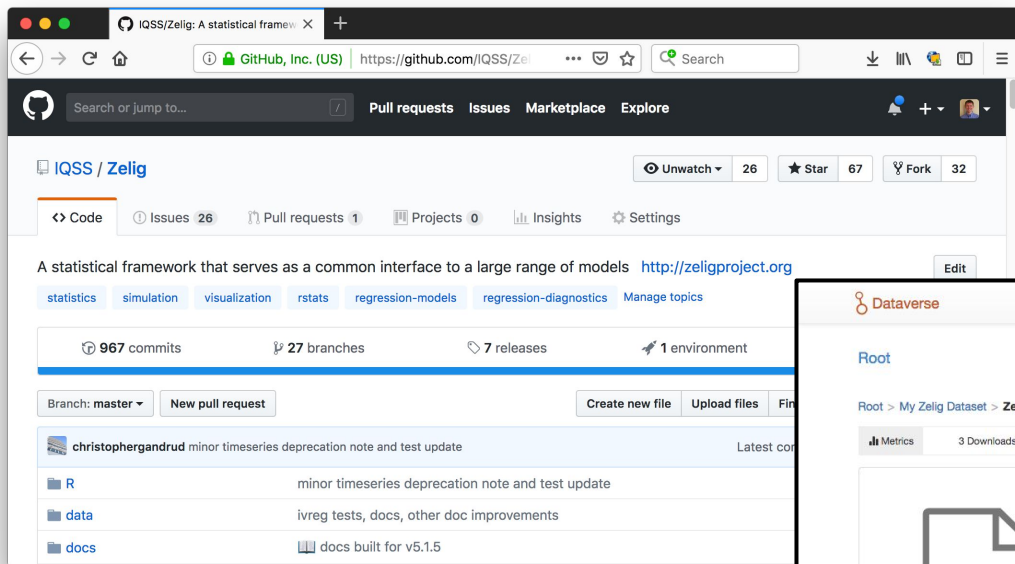
Figure4.R
Jan 18, 2019 - Replication Data for: When Should We Use Unit Fixed Effects Regression Models for Causal Inference with Longitudinal Data?
R Syntax - 8.1 KB - MD5: cb92b3afa8ee636fd454bdf917c890
[Code](#)

Figure5.R
Jan 18, 2019 - Replication Data for: When Should We Use Unit Fixed Effects Regression Models for Causal Inference with Longitudinal Data?
R Syntax - 3.7 KB - MD5: 9b0a935e31e173903faaf6294a937445
[Code](#)

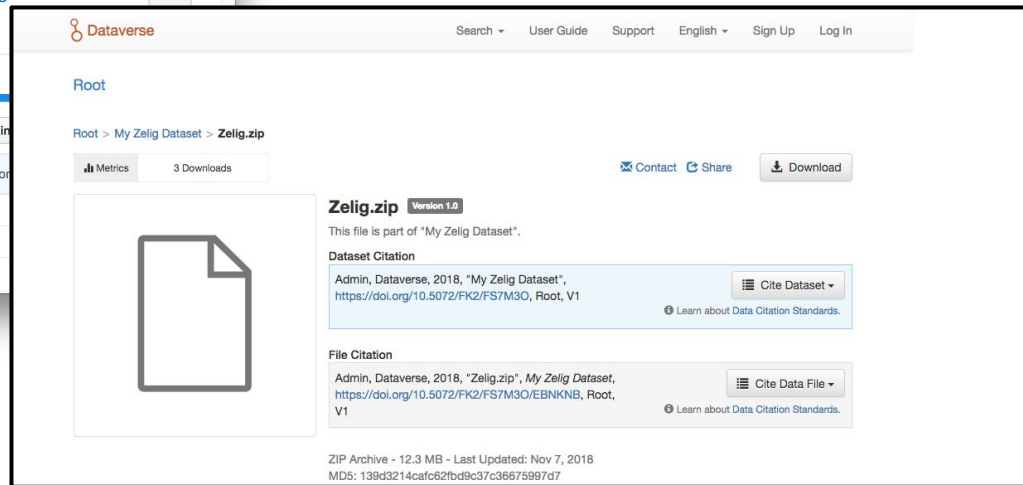
Table2.R
Jan 18, 2019 - Replication Data for: When Should We Use Unit Fixed Effects Regression Models for Causal Inference with Longitudinal Data?
R Syntax - 3.1 KB - MD5: 43d3a5072d0229a094c16254cb026bb0
[Code](#)

[Feedback](#)

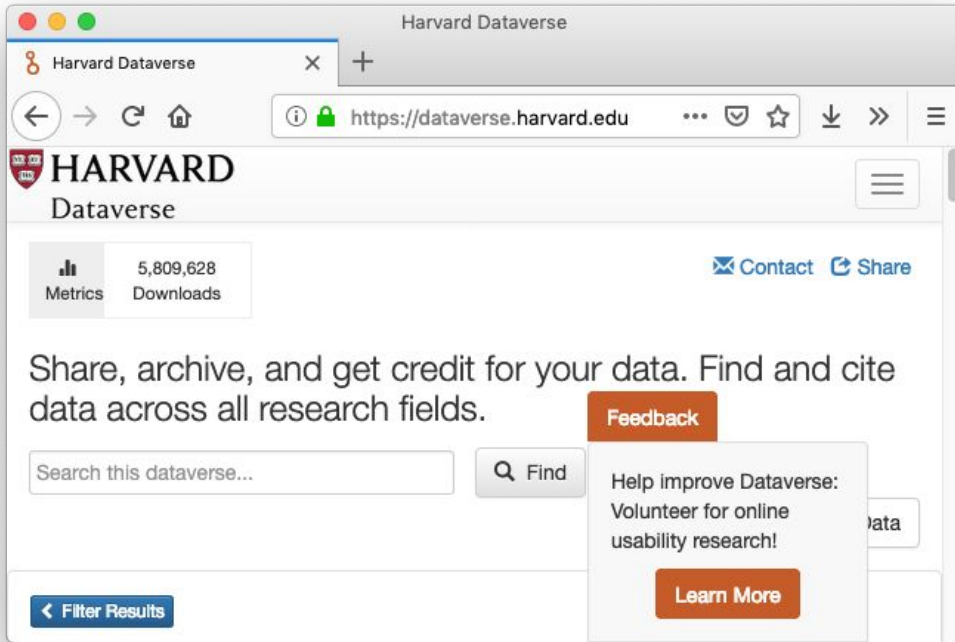
Import from GitHub prototype



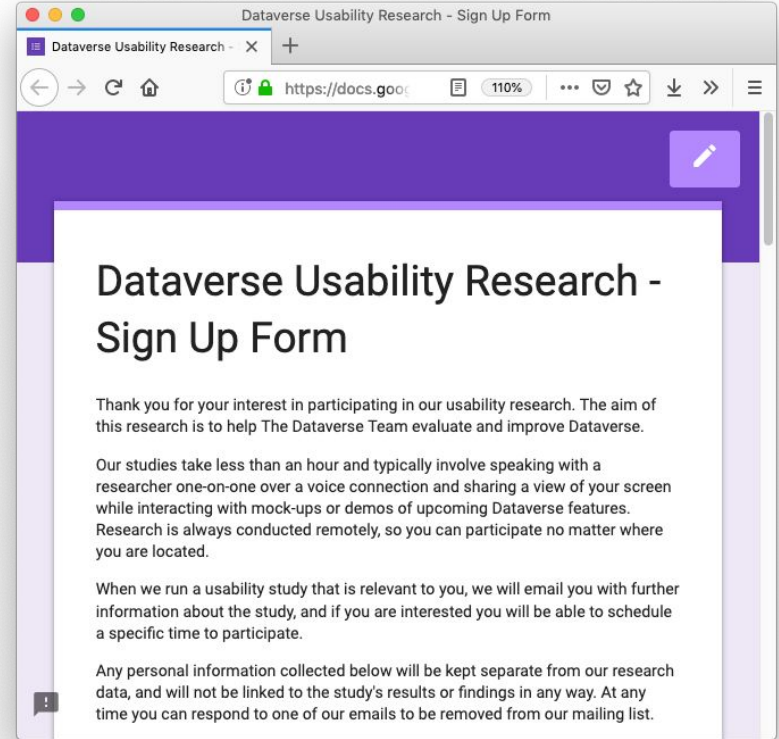
<https://github.com/IQSS/dataverse/issues/2739>



Usability testing



<https://dataverse.harvard.edu>



Software metadata

Files **Metadata** Terms Versions

Search this dataset...

1 to 10 of 29 Files

<input type="checkbox"/>	Download
<input type="checkbox"/>	Download
<input type="checkbox"/>	Download
<input type="checkbox"/>	Download
<input type="checkbox"/>	Download

1_Readme_and_DataFormats.zip
ZIP Archive - 64.1 KB - Sep 19, 2017 - 49 Downloads
MD5: a4de02435e7eec49c422e1d45cb60758
Contains the main readme file, the data formats file, which describes the different types of data found within the other files, and a metadata html file documenting the software included with the project.
Documentation

2_Fishtracker.zip
ZIP Archive - 15.2 MB - Sep 19, 2017 - 46 Downloads
MD5: bb5a1d3d3e32461796e4ee0c345472b9
Codebase for the fish tracking algorithm described in the paper. (Readme enclosed)
Code

3_PaperFigs.zip
ZIP Archive - 83.3 KB - Sep 19, 2017 - 46 Downloads
MD5: fed6b315f84284945bca1cbb7070e643
Code used to generate the data and results figures from the paper. (Readme enclosed)
Code

4_AnalysisCommon.zip
ZIP Archive - 388.9 KB - Sep 19, 2017 - 43 Downloads
MD5: 87867b1a13af3ba280fc3499bb8a1765
Common code for analyzing laboratory data. analyzeData.m from each lab data folder (File 6-02, 7-02, 8-02 and 9-02) uses code from this folder to initially process the raw data. (Readme enclosed)
Code

<http://datacuratornetwork.org>

Data and software associated with ...

https://archive.data.jhu.edu/dataset.xhtml?persistentId=... Search

JOHNS HOPKINS LIBRARIES **Johns Hopkins Data Services**

Dataverse Search User Guide Support Log In

Citation Metadata

Software Description

Software Title	Software associated with publication "High-resolution behavioral mapping of electric fishes in Amazonian habitats"
License	MIT license (MIT)
Description	'Fishtracker' is the two-step algorithm described in Madhav, Jayakumar et. al, 2017 to track multiple freely moving weakly electric fish using measurements from a grid of electrodes in a known configuration (see '2_Fishtracker'). Also included is code used to generate the data and results figures from the Madhav, Jayakumar et. al , 2017 (see '3_PaperFigs') and code for analyzing laboratory data (see '4_AnalysisCommon').
Artifact Type	Source Code
Date Published	2017-04
Code Repository Link	https://github.com/manusmad/fishtracker
Programming Language	MATLAB
Function	Data collection / organization; Visualization
Interaction Method	Graphical; Programmatic
Software Contributors	Madhav, Manu Developer Jayakumar, Ravikrishnan Developer
Dependencies	Software Package or Platform - MATLAB v. R2016a Library / Module - CircStat2012a, Hungarian, MagnetGInput, SONlib, addpath_recurse, datastructure, dispense, distinguishable_colors, export_fig, findjob, ginputax, matson, ndnanfilter, parfor_progress, plot_ellipse, progressbar, savefast, serialization, subplotplot, Refer to Readme in 2_Fishtracker for more information.
Other related software	Interoperable Software - File 3_PaperFigs (download available from this project): code used to generate the data and results figures from the Madhav, Jayakumar et. al , 2017. Interoperable Software - File 4_AnalysisCommon (download available from this project): code for analyzing laboratory data. analyzeData.m from each lab data folder included in this project uses code from this folder to initially process the raw data.

Copyright © 2019 Johns Hopkins University

Powered by **The Dataverse Project**

CodeMeta

Supported Content Types

DataCite supports a number of metadata content types:

Format	Content Type
RDF XML	application/rdf+xml
RDF Turtle	text/turtle
Citeproc JSON	application/vnd.citationstyles.csl+json
Schema.org in JSON-LD	application/vnd.schemaorg.ld+json
Codemeta	application/vnd.codemeta.ld+json
Formatted text citation	text/x-bibliography
RIS	application/x-research-info-systems
BibTeX	application/x-bibtex
DataCite XML	application/vnd.datacite.datacite+xml
DataCite JSON	application/vnd.datacite.datacite+json

<https://support.datacite.org/docs/datacite-content-resolver>



DataCite

FIND, ACCESS, AND REUSE DATA

Enhanced software preservation now available in CaltechDATA!

Friday, March 09, 2018

CaltechDATA has supported automatic preservation of GitHub software repositories since launch, so anyone at Caltech can get a DOI (permanent identifier) for their software project and have Caltech Library handle long term preservation. However, most GitHub repositories do not include clear metadata such as authors, affiliations, or ORCID identifiers. CaltechDATA now supports [CodeMeta](#), a new standard format for software metadata. By including a [codemeta.json](#) file in your GitHub repo, your full author list, keywords, and license will be listed in CaltechDATA and registered with your DOI.

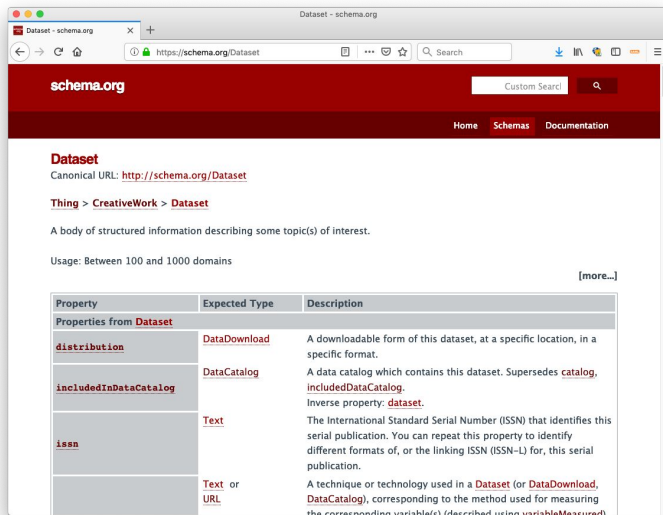
This improvement is powered by [ames](#), a Python package for automating metadata changes developed at Caltech Library. Every 5 minutes, ames harvests all the GitHub-created records in CaltechDATA and stores them using [dataset](#) (our lightweight data storage package). These records are then analyzed for codemeta.json files. If a CodeMeta file is found, the relevant metadata is extracted and added to the CaltechDATA record and DOI. We currently support authors, keywords, and license fields - but more will be added as a community of practice develops. We're also exploring better ways to generate CodeMeta files as part of the software release process.

<https://www.library.caltech.edu/news/enhanced-software-preservation-now-available-caltechdata>

Schema.org JSON-LD

Google Dataset Search Beta

Search for Datasets



Dataset
Canonical URL: <http://schema.org/Dataset>

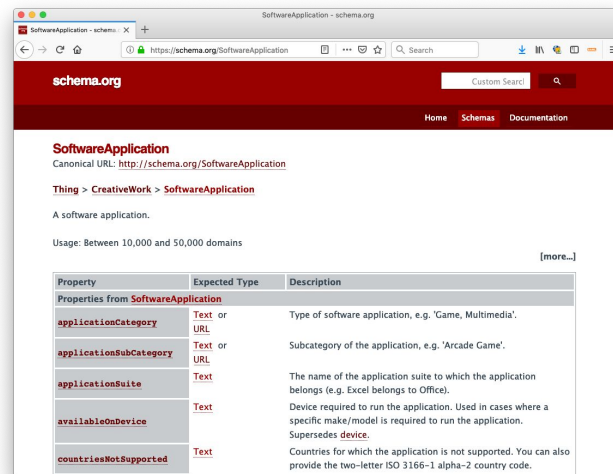
Thing > CreativeWork > Dataset

A body of structured information describing some topic(s) of interest.

Usage: Between 100 and 1000 domains [more..]

Property	Expected Type	Description
Properties from Dataset		
distribution	DataDownload	A downloadable form of this dataset, at a specific location, in a specific format.
includedInDataCatalog	DataCatalog	A data catalog which contains this dataset. Supersedes catalog , includedDataCatalog . Inverse property: dataset .
issn	Text	The International Standard Serial Number (ISSN) that identifies this serial publication. You can repeat this property to identify different formats of, or the linking ISSN (ISSN-L) for, this serial publication.
	Text or URL	A technique or technology used in a Dataset (or DataDownload , DataCatalog), corresponding to the method used for measuring the corresponding variable(s) described using variableMeasured.

<https://schema.org/Dataset>



SoftwareApplication
Canonical URL: <http://schema.org/SoftwareApplication>

Thing > CreativeWork > SoftwareApplication

A software application.

Usage: Between 10,000 and 50,000 domains [more..]

Property	Expected Type	Description
Properties from SoftwareApplication		
applicationCategory	Text or URL	Type of software application, e.g. 'Game, Multimedia'.
applicationSubCategory	Text or URL	Subcategory of the application, e.g. 'Arcade Game'.
applicationSuite	Text	The name of the application suite to which the application belongs (e.g. Excel belongs to Office).
availableOnDevice	Text	Device required to run the application. Used in cases where a specific make/model is required to run the application. Supersedes device .
countriesNotSupported	Text	Countries for which the application is not supported. You can also provide the two-letter ISO 3166-1 alpha-2 country code.

<https://schema.org/SoftwareApplication>

Software citation

The screenshot shows the website for the Software Citation Implementation Working Group (FORCE11). The page features a navigation menu with options like 'ABOUT', 'COMMUNITY', 'GROUPS', 'RESOURCES', 'NEWS + BLOGS', 'EVENTS', 'PUBLICATIONS', and 'MEDIA'. A 'GROUP MENU' sidebar lists 'Group Home', 'Members', 'Links + Files', 'Google Forum', and 'Subscribe to group'. The main content area is titled 'SOFTWARE CITATION IMPLEMENTATION WORKING GROUP' and includes a 'Description' section with a list of goals: 1. endorse the principles, 2. develop sets of guidelines for implementing the principles, 3. help implement the principles, and 4. test specific implementations of the principles. It also lists group leaders: Neil Chue Hong, Martin Fenner, and Daniel S. Katz, and provides a group email address: softwarecitationimplementationwg@force11.org.



Software Citation Principles

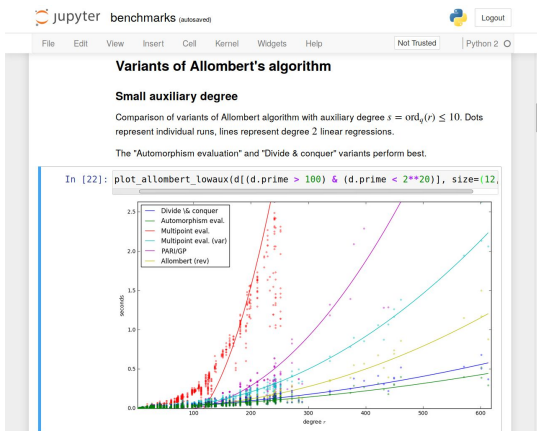
<https://doi.org/10.7717/peerj-cs.86>

- **Importance:** Software should be considered a legitimate and citable product of research...
- **Credit and attribution:** ...normative, legal attribution to all contributors to the software...
- **Unique identification:** ...identification that is machine actionable, globally unique, interoperable, and recognized...
- **Persistence:** Unique identifiers and metadata describing the software and its disposition should persist...
- **Accessibility:** ...[access to the software itself](#) and to its associated metadata...
- **Specificity:** ...identification of, and access to, the specific version of software...

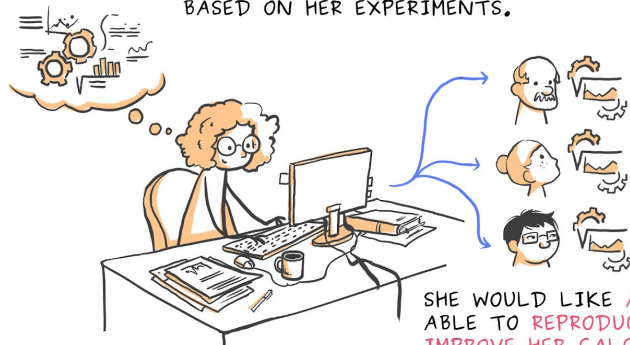
https://slides.com/dbouquin/datafest2019_citesoftware

<https://www.force11.org/group/software-citation-implementation-working-group>

Reproducibility



JANE HAS WRITTEN A PAPER BASED ON HER EXPERIMENTS.



SHE WOULD LIKE ANYONE TO BE ABLE TO REPRODUCE, CHECK, AND IMPROVE HER CALCULATIONS

STEP 1

SHE DESCRIBES THE EXPERIMENTS AS A Jupyter **NOTEBOOK**, MIXING:

PROSE
CODE &
VISUALIZATION

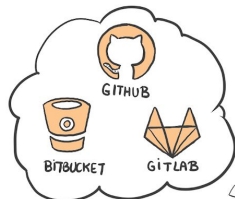
AND RESOURCES:

SOURCE CODE,
DATA,
MEDIA...

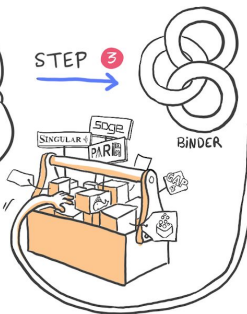
STEP 2

SHE PUBLISHES THEM ON A PUBLICLY HOSTED REPOSITORY

SHE MAKES THAT REPOSITORY **BINDER-READY** BY DESCRIBING THE SOFTWARE REQUIRED TO RUN THE NOTEBOOK



STEP 3



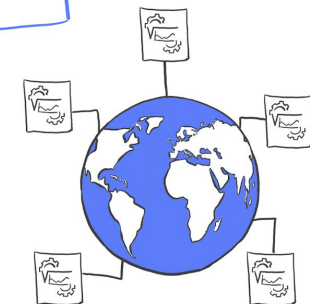
CONFIGURATION ✓



NOTEBOOK ✓

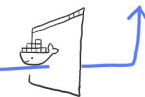


RESOURCES ✓



EVERYONE CAN NOW RUN AND REPRODUCE HER COMPUTATIONS

STEP 4



Code Ocean (Sloan grant)



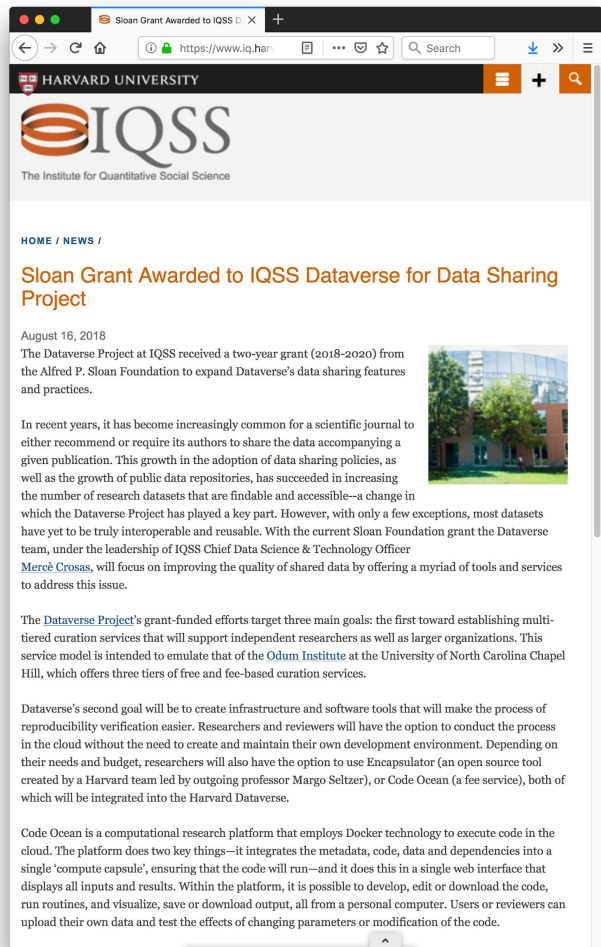
<https://twitter.com/mercecrosas/status/1072899669074821122>



CODE OCEAN

Code Ocean is a
cloud-based
computational
reproducibility
platform

<https://www.iq.harvard.edu/news/sloan-grant-awarded-iqss-dataverse-data-sharing-project>



Code Ocean



CODE OCEAN

Rutgers University Workshop Example - Code | Code Ocean

https://codeocean.com/capsule/cdf2e8f4-ed9e-4716-a890-87346219e1...

Rutgers University Workshop Example

Capsule File Edit View Tabs Settings Help

Viewing

Switch to Old Editor Sign up

Run (run.sh)

or, launch interactive session

Environment & Dependencies

Run 2128428 0:00:02
April Clyburne-Sherin | 17 days ago
output 239 B

Run 563096 0:00:08
April Clyburne-Sherin | about 1 month ago

Run 5426463 0:00:09
April Clyburne-Sherin | 3 months ago

```
1 #!/bin/bash
2
3 python fig1_happiness_of_individuals.py
4 python fig2_distribution_of_happiness.py
5
6 #pip freeze > ../results/requirements.txt
7 conda list --export > ../results/requirements.txt
8
9 jupyter nbconvert \
10 --ExecutePreprocessor.allow_errors=True \
11 --ExecutePreprocessor.timeout=1 \
12 --FilesWriter.build_directory=../results \
13 --execute candy_trade.ipynb
14
15
16
17
18
19
```

<http://bit.ly/harvard-oa-week>

https://github.com/aprilcs/candy_trade

<https://github.com/IQSS/dataverse/issues/5028>

Harvard University Workshop Example (copy) - Code | Code Ocean

https://codeocean.com/capsule/b4d3288f-8df4-4a48-a5...

Harvard University Workshop Example (copy)

Capsule File Edit View Tabs Settings Help

Switch to Old Editor

Run (run.sh)

or, launch interactive session

Environment & Dependencies

Publish capsule & results

Run 481696 0:00:08
Philip Durbin | about 1 month ago
output 326 B
candy_trade.html 309.81 KB
fig1_happiness_of_individuals.png 60.11 KB
fig2_distribution_of_happiness.png 20.37 KB
requirements.txt 1.28 KB

Happiness of individuals with candy selection vs. number of candy trades

Happiness Rating

Number of trades

Individual	0.00	0.25	0.50	0.75	1.00	1.25	1.50	1.75	2.00
Amelia	6.0	6.5	7.0	7.5	8.0	8.5	9.0	9.5	10.0
Daniel	6.0	6.5	7.0	7.5	8.0	8.5	9.0	9.5	10.0
Daniel_L	6.0	6.5	7.0	7.5	8.0	8.5	9.0	9.5	10.0
Dianne	6.0	6.5	7.0	7.5	8.0	8.5	9.0	9.5	10.0
German	6.0	6.5	7.0	7.5	8.0	8.5	9.0	9.5	10.0
Grace	6.0	6.5	7.0	7.5	8.0	8.5	9.0	9.5	10.0
Henrik	6.0	6.5	7.0	7.5	8.0	8.5	9.0	9.5	10.0
Hester	6.0	6.5	7.0	7.5	8.0	8.5	9.0	9.5	10.0
Julien	6.0	6.5	7.0	7.5	8.0	8.5	9.0	9.5	10.0
Katy	6.0	6.5	7.0	7.5	8.0	8.5	9.0	9.5	10.0
Sunsiree	6.0	6.5	7.0	7.5	8.0	8.5	9.0	9.5	10.0

Replication Datasets



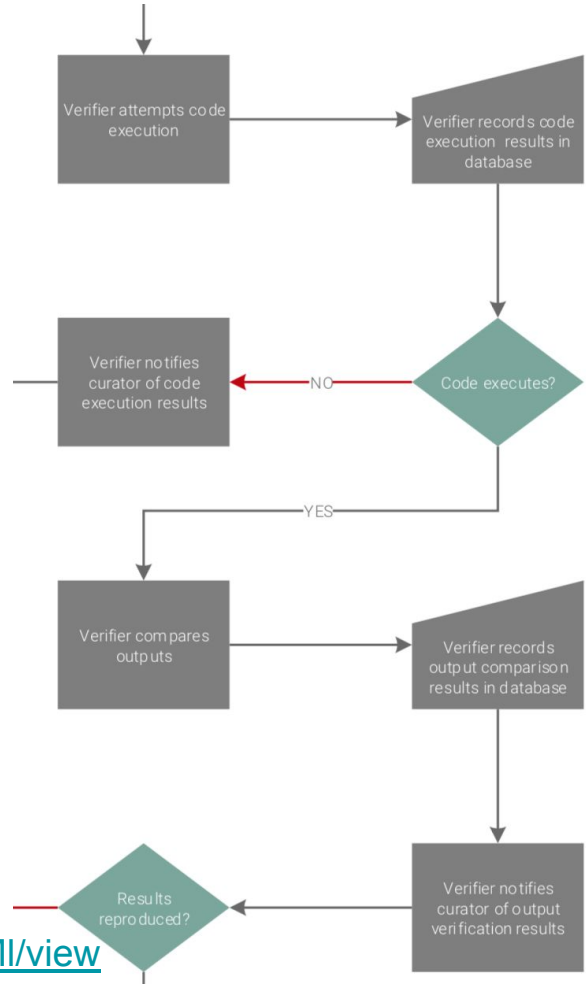
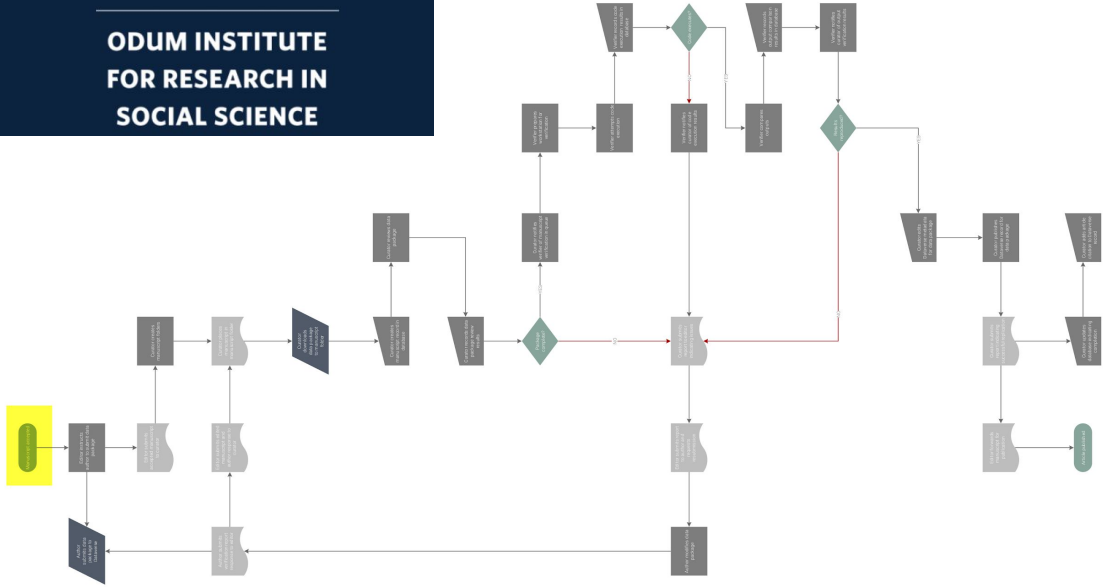
The screenshot shows the Harvard Dataverse interface. At the top, it says "HARVARD Dataverse" with navigation links for "About", "User Guide", "Support", "Sign Up", and "Log In". Below this is a banner for the "AMERICAN JOURNAL of POLITICAL SCIENCE" (AJPS) Dataserve, hosted by Michigan State University. The main content area displays the title "Replication Data for: No Harm in Checking: Using Factual Manipulation Checks to Assess Attentiveness in Experiments" (Version 1.0) by Kane, John V. (2018). The description states: "Manipulation checks are often advisable in experimental studies, yet they rarely appear in practice. This lack of usage may stem from fears of distorting treatment effects and uncertainty regarding which type to use (e.g., instructional manipulation checks [IMCs] or assessments of whether stimuli alter a latent independent variable of interest). Here, we first categorize the main variants and argue that factual manipulation checks (FMCs)—i.e., objective questions about key elements of the experiment—can identify individual-level attentiveness to experimental information and, as a consequence, better enable researchers to diagnose experimental findings. We then find, through four replication studies, little evidence that FMC placement affects treatment effects, and that placing FMCs immediately post-outcome does not attenuate FMC passage rates. Additionally, FMC and IMC passage rates are only weakly related, suggesting that each technique identifies different sets of attentive subjects. Thus, unlike other methods, FMCs can confirm attentiveness to experimental protocols. (2018-06-25)". The subject is "Social Sciences" and the keyword is "Manipulation check, Screener". A note mentions that the dataset underwent an independent verification process at the Odum Institute for Research in Social Science at UNC-CH. At the bottom, there are "Open Data" and "Open Materials" badges and a "Feedback" button.



"This dataset underwent an independent verification process that **replicated the tables and figures in the primary article**. For the supplementary materials, verification was performed solely for the **successful execution of code**. The verification process was carried out by the Odum Institute for Research in Social Science at the University of North Carolina at Chapel Hill."

6000 replication datasets in Harvard Dataverse: <https://dataverse.org/best-practices/replication-dataset>

Verification workflow



<https://drive.google.com/file/d/1W4gdckYXh9fANLdTloTydJRGtb2E9PMI/view>

Open Science Infrastructure Working Group

Description

The proposed working group will bring together representatives from key open science infrastructure projects related to computational reproducibility and publishing of composite research objects (e.g., code, data, and environment) to improve interoperability and to coordinate collaborative development as needed. A central goal of the project is to maximize the utility of developed tools while minimizing duplication of effort. The working group may also expand to cover technical approaches to tracking provenance.

Meeting notes

- [Meeting notes](#)
- Related Github issues:
 - <https://github.com/whole-tale/whole-tale/issues/43>
 - <https://github.com/whole-tale/whole-tale/issues/50>

Outputs

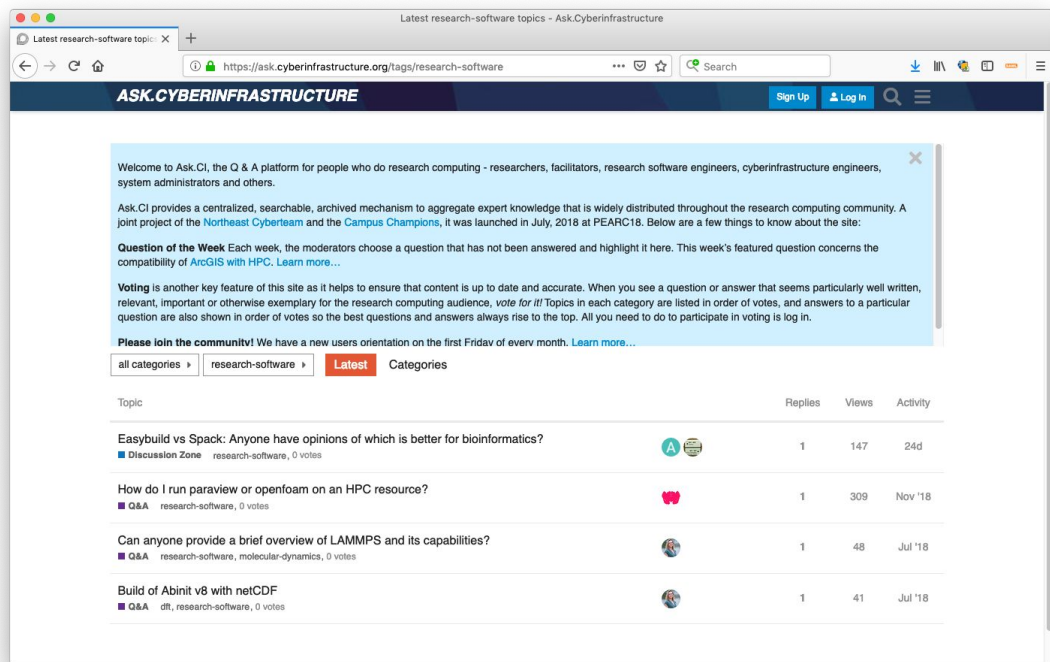
- Recommendations/best practices for publishing composite objects to research repositories for re-execution.
- Define points of interoperability across projects
- Identify areas where WT project can leverage and contribute to existing community initiatives (e.g., repo2docker)

https://github.com/whole-tale/whole-tale/tree/master/working_groups/open-infrastructure

2019 Dataverse Community Meeting: June 19-21



Ask.CI research software



The screenshot shows a web browser window displaying the Ask.CI website. The page title is "Latest research-software topics - Ask.Cyberinfrastructure". The URL in the address bar is "https://ask.cyberinfrastructure.org/tags/research-software". The page features a dark blue header with the "ASK.CYBERINFRASTRUCTURE" logo, "Sign Up", and "Log In" buttons. A light blue banner contains a welcome message and information about the site's purpose and features, including a "Question of the Week" and a "Please join the community!" notice. Below the banner, there are navigation tabs for "all categories", "research-software", "Latest", and "Categories". The main content area displays a list of questions with columns for "Topic", "Replies", "Views", and "Activity".

Welcome to Ask.CI, the Q & A platform for people who do research computing - researchers, facilitators, research software engineers, cyberinfrastructure engineers, system administrators and others.

Ask.CI provides a centralized, searchable, archived mechanism to aggregate expert knowledge that is widely distributed throughout the research computing community. A joint project of the [Northeast Cyberteam](#) and the [Campus Champions](#), it was launched in July, 2018 at PEARC18. Below are a few things to know about the site:

Question of the Week Each week, the moderators choose a question that has not been answered and highlight it here. This week's featured question concerns the compatibility of [ArcGIS with HPC](#). [Learn more...](#)

Voting is another key feature of this site as it helps to ensure that content is up to date and accurate. When you see a question or answer that seems particularly well written, relevant, important or otherwise exemplary for the research computing audience, *vote for it!* Topics in each category are listed in order of votes, and answers to a particular question are also shown in order of votes so the best questions and answers always rise to the top. All you need to do to participate in voting is log in.

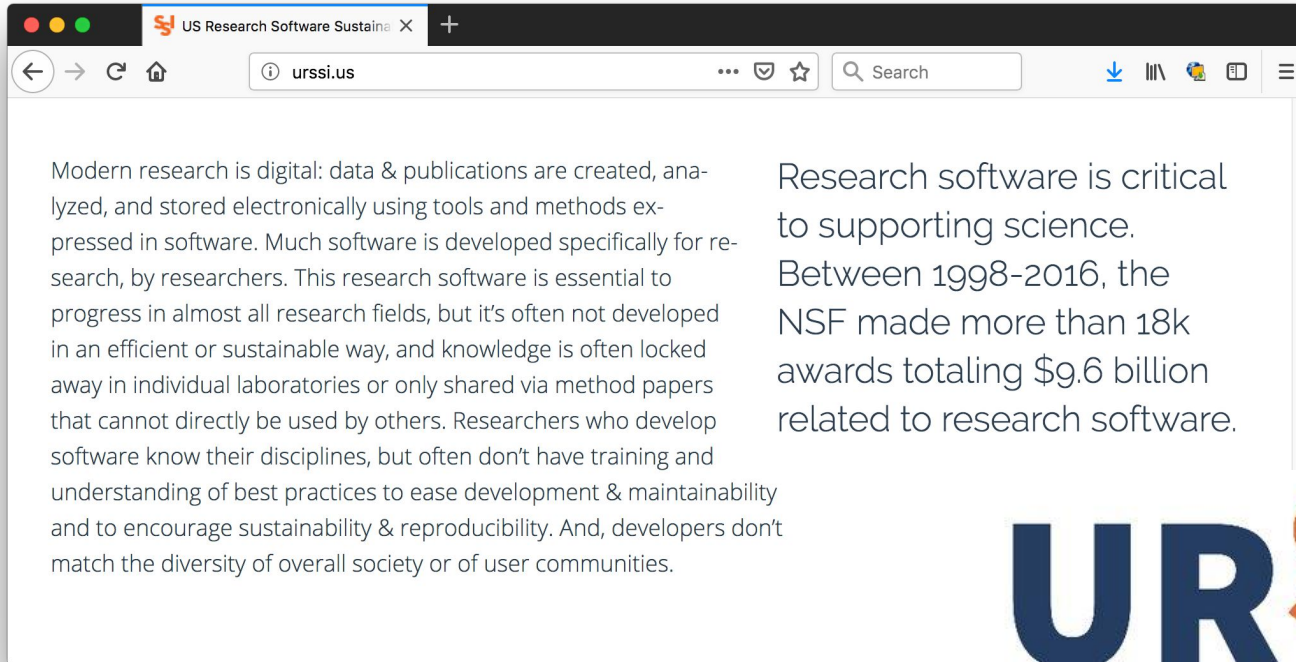
Please join the community! We have a new users orientation on the first Friday of every month. [Learn more...](#)

all categories > research-software > **Latest** Categories

Topic	Replies	Views	Activity
Easybuild vs Spack: Anyone have opinions of which is better for bioinformatics? ■ Discussion Zone research-software, 0 votes	1	147	24d
How do I run paraview or openfoam on an HPC resource? ■ Q&A research-software, 0 votes	1	309	Nov '18
Can anyone provide a brief overview of LAMMPS and its capabilities? ■ Q&A research-software, molecular-dynamics, 0 votes	1	48	Jul '18
Build of Abinit v8 with netCDF ■ Q&A dtl, research-software, 0 votes	1	41	Jul '18

<https://ask.cyberinfrastructure.org/tags/research-software>

US Research Software Sustainability Institute



The image shows a browser window with the URL `urssi.us`. The page content is split into two columns. The left column contains a paragraph of text, and the right column contains a shorter paragraph. The browser's address bar and navigation icons are visible at the top.

Modern research is digital: data & publications are created, analyzed, and stored electronically using tools and methods expressed in software. Much software is developed specifically for research, by researchers. This research software is essential to progress in almost all research fields, but it's often not developed in an efficient or sustainable way, and knowledge is often locked away in individual laboratories or only shared via method papers that cannot directly be used by others. Researchers who develop software know their disciplines, but often don't have training and understanding of best practices to ease development & maintainability and to encourage sustainability & reproducibility. And, developers don't match the diversity of overall society or of user communities.

Research software is critical to supporting science. Between 1998-2016, the NSF made more than 18k awards totaling \$9.6 billion related to research software.



<https://discuss.urssi.us/t/recent-talks-on-research-software/30>

Thank you!

[@philipdurbin](https://twitter.com/philipdurbin)

[@dataverseorg](https://twitter.com/dataverseorg)

support@dataverse.org



Open source research data repository software



Researchers

Enjoy full control over your data. Receive *web visibility*, *academic credit*, and *increased citation counts*. A personal dataverse 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?](#)



Journals

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 Dataverse as part of your journal data policy or list of repository recommendations. [Want to find out more about journal dataverses?](#)



Institutions

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

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 Dataverse. [Want to contribute?](#)

<https://dataverse.org>

My questions for you

- Are potential and existing code depositors to Dataverse comfortable with Dataverse's current versioning convention being applied to code? Why or why not?
- Do current or potential code depositors want to apply the version number from GitHub to the code they publish in Dataverse? Why or why not?
- Do users of code in Dataverse expect a relationship between the version in Dataverse and the releases in GitHub? Why or why not?