

Failure to Reproduce: The Replication Crisis in Research – Can Librarians Help?

Andrée Rathemacher, Amanda Izenstark, & Harrison Dekker
University of Rhode Island Libraries

May 4, 2018
ACRL New England Annual Conference

1.The Reproducibility Crisis

2.Reproducible Workflows

3.Introduction to the
Open Science Framework

The Reproducibility Crisis

“It can be proven that most claimed research findings are false.”

– John P. A. Ioannidis, 2005

“Reproducibility crisis” (aka “replication crisis”)

“A methodological crisis in science in which scientists have found that the results of many scientific experiments are difficult or impossible to replicate on subsequent investigation, either by independent researchers or by the original researchers themselves.”

– Wikipedia

Psychology



91.5% of all published studies in psychology found positive results.

“[EEG Experiment](#)”
from Dr. Hirt’s
Psychology Lab,
Indiana University

Economics

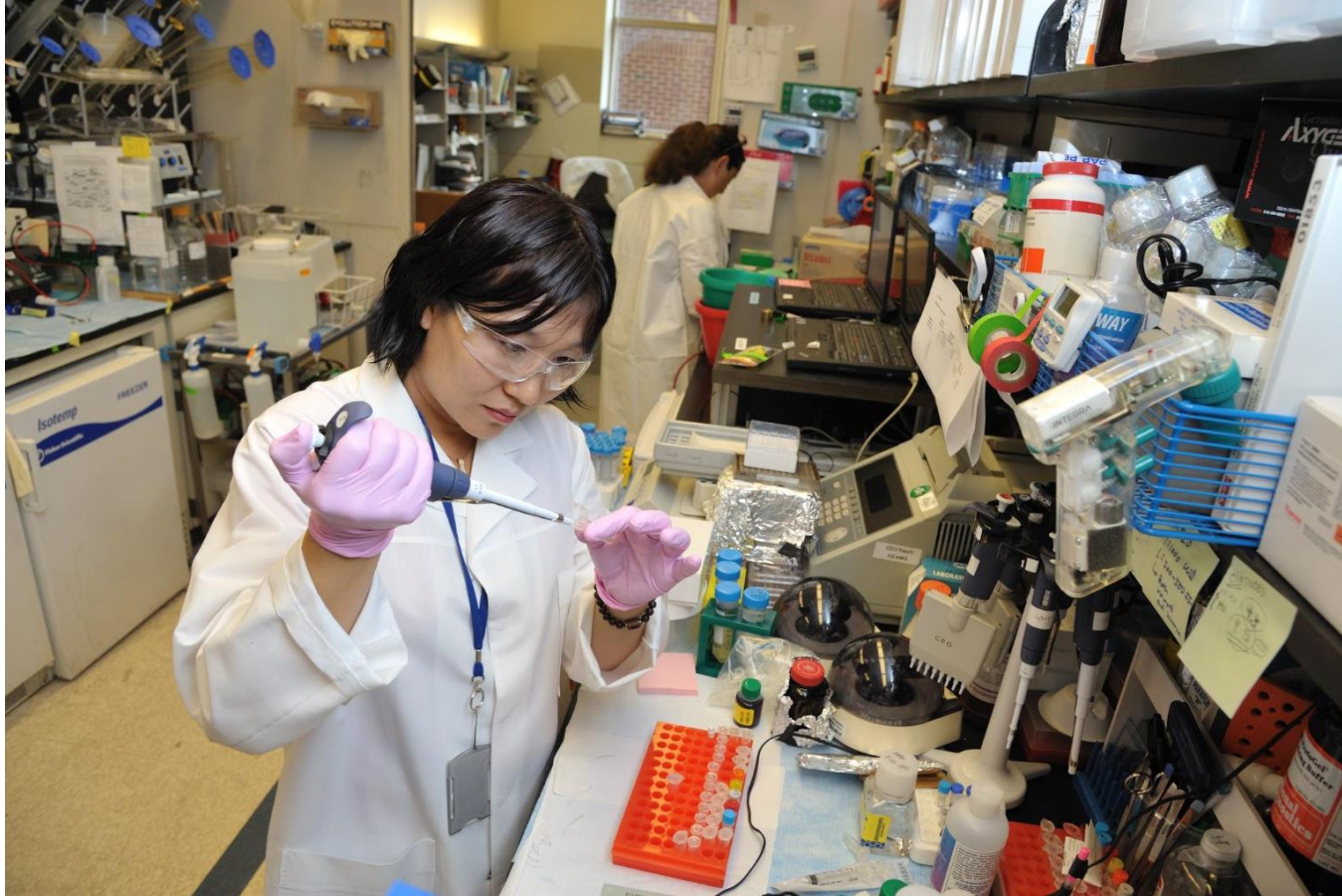


“...We assert that economics research is usually not replicable.”

– Andrew C. Chang and Phillip Li, 2015

“[Homeless man in Vancouver](#)” by Jay Black is licensed under [CC BY-SA 2.0](#).

Biomedical research



“[The NIAMS Cartilage Biology and Orthopaedics Branch](#)” by [NIH Image Gallery](#) is licensed under [CC BY-NC 2.0](#).

Why? “File-drawer problem”

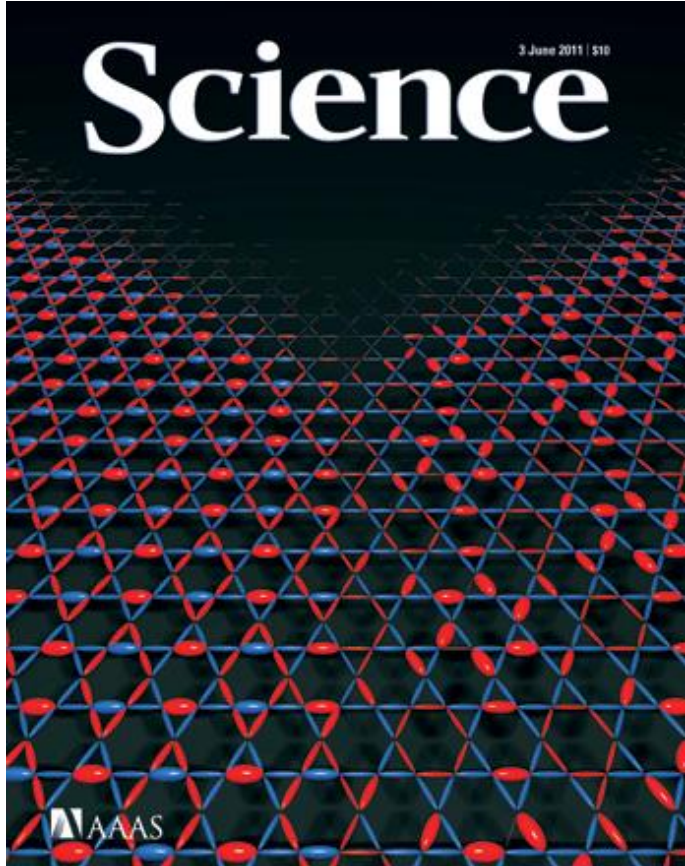


Researchers do not bother to write up experiments with negative / null results or the results of replication studies.

Instead of submitting them to journals, they file them away.

“Filing” by [Jeff Youngstrom](#) is licensed under [CC BY-NC 2.0](#).

Why? *Publication bias*

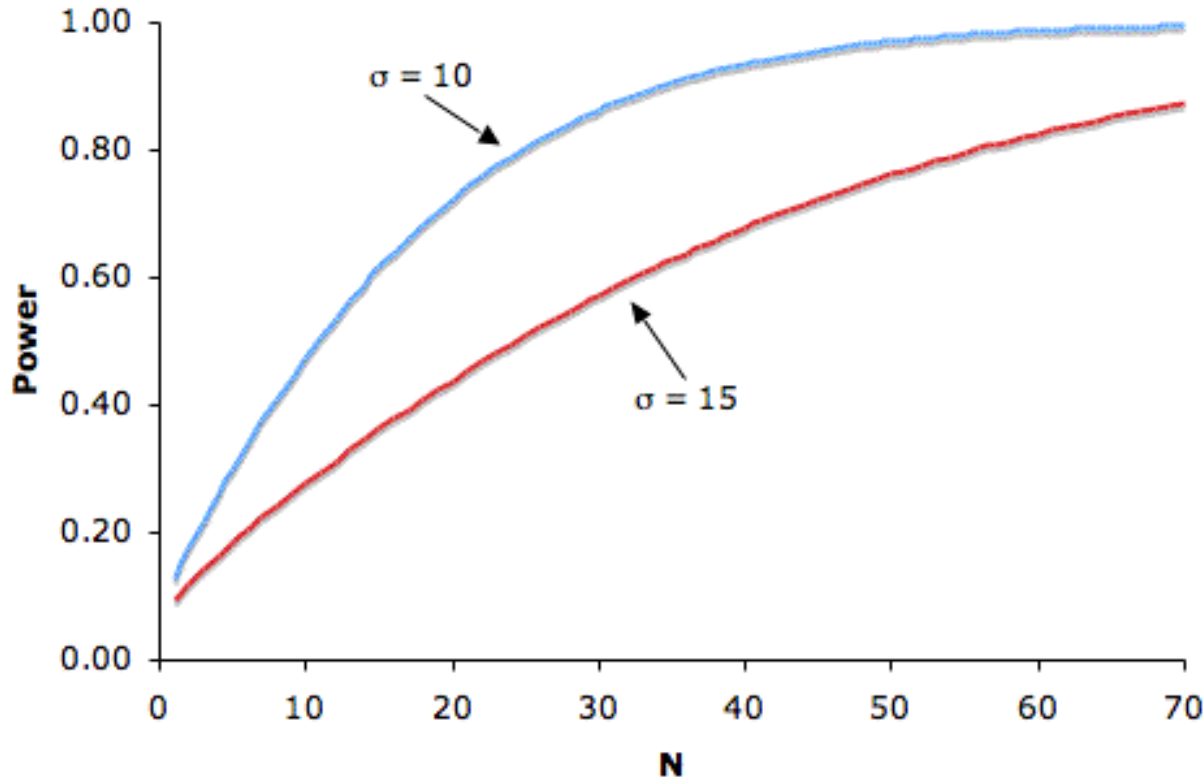


“...the small proportion of results chosen for publication are unrepresentative of scientists’ repeated samplings of the real world.”

– Neal S. Young, John P. A. Ioannidis, and Omar Al-Ubaydli, 2008

[Cover of Science v. 332, no. 6034](#) by the American Association for the Advancement of Science. Image by Stephen R. White.

Why? *Bad experimental design & analysis*



“If you torture the data long enough, it will confess.”

– Ronald Coase, recipient of the 1991 Nobel Prize in Economics

[“The Relationship Between Sample Size and Power”](#) by [Online Statistics Education: A Multimedia Course of Study](#) is in the public domain.

Why? *Incentive structure*



“Today I wouldn’t get an academic job. It’s as simple as that. I don’t think I would be regarded as productive enough.”

– Peter Higgs, 2013 (winner of the 2013 Nobel Prize in Physics)

“[Prof. Meyerson in his funky Stanford gown](#)” by [Anna Majkowska](#) is licensed under [CC BY 2.0](#).

What about peer review?



“We need to get away from the notion, proven wrong on a daily basis, that peer review of any kind at any journal means that a work of science is correct.”

– Michael Eisen, 2014

“[Peer Review Monster](#)” by [Gideon Burton](#) is licensed under [CC BY-SA 2.0](#).

Reproducible Workflows

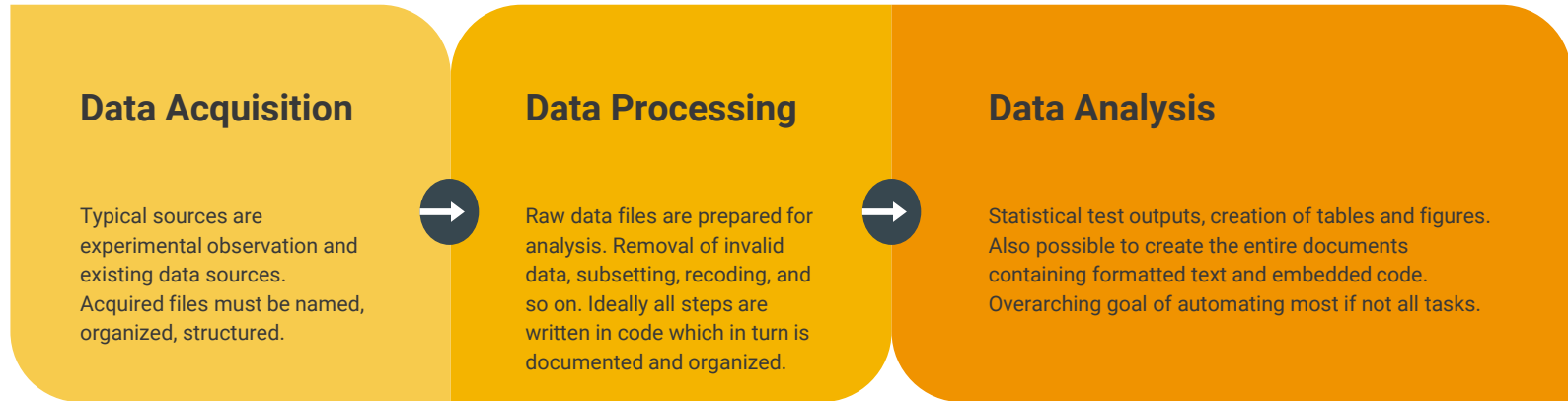
Scholarship or advertising?



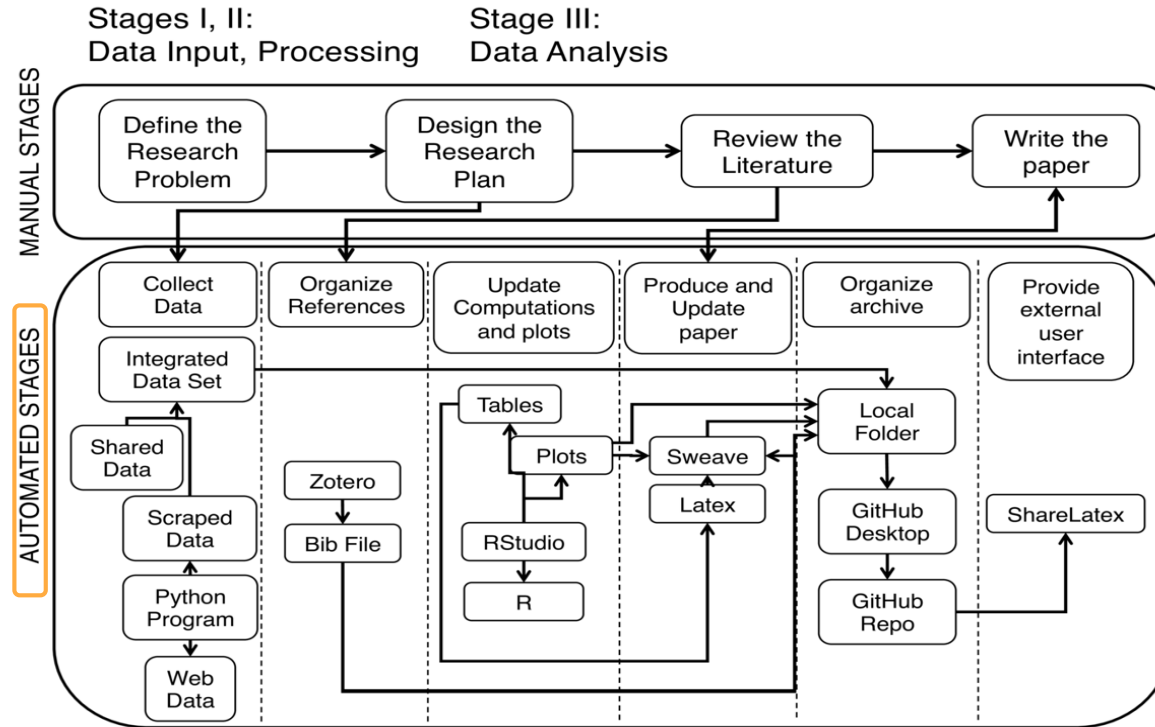
An article about computational science in a scientific publication is not the scholarship itself, it is merely advertising of the scholarship. The actual scholarship is the complete software development environment and the complete set of instructions which generate the figures.

- Jonathan Buckheit and David Donoho, 1995

General workflow model



Detailed workflow example



Workflow skills and tools

Skill type	Description	Tools
Literate computing	Enable writing self-contained documents combining text and code	Rstudio : Markdown : LaTeX : Jupyter
Version control	Track file changes over time. Revert to earlier versions. Branch/fork	Git : GitHub : BitBucket
Tracking provenance	Capture complex workflows involving multiple research objects/tools	VisTrails : Kepler : Taverna
Automation	Automate workflows using time-tested and ubiquitous command line tools	Unix command line : shell scripts : make
Virtual environments	Capture complex computation environments and configurations	VirtualBox : VMWare : Docker

Learning incentives



The first step to making science reproducible is to build good habits. Your most important collaborator is your future self. It's important to make a workflow that you can use time and time again, and even pass on to others in such a way that you don't have to be there to walk them through it.

Culich, 2014

More information

Training:

Data/Software Carpentry -- <https://carpentries.org/>

Library Carpentry -- <https://librarycarpentry.github.io/>

Case Studies:

Kitzes, J., Turek, D., & Deniz, F. (Eds.). (2018). *The Practice of Reproducible Research: Case Studies and Lessons from the Data-Intensive Sciences*. Oakland, CA: University of California Press.

(A free [pre-print edition](#) is available)

Teaching materials:

Project TIER -- <https://www.projecttier.org/>

ROpenSci -- <https://ropensci.github.io/reproducibility-guide/>

BITSS -- <https://www.bitss.org/resources/>

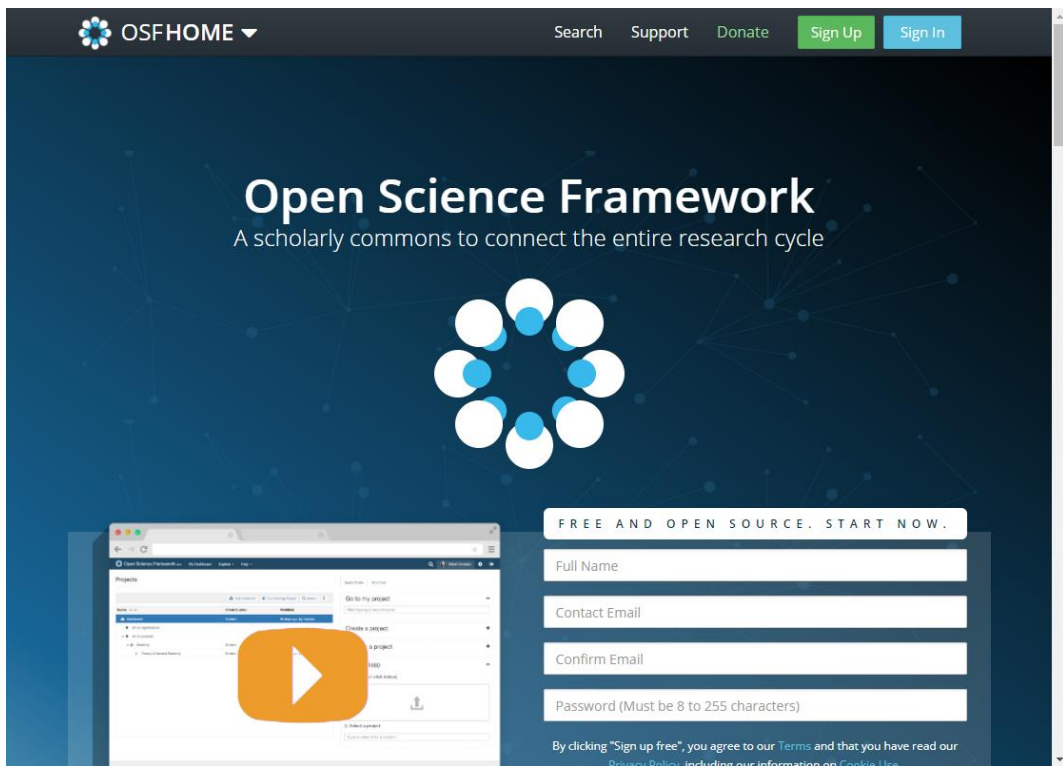
Introduction to The Open Science Framework

Why the Open Science Framework?

Project of the
Center for Open Science,
a nonprofit based in
Charlottesville, VA

Funded by a variety of
grants and sponsors,
including DARPA, the
NSF, NIH, and others.

<https://osf.io/>



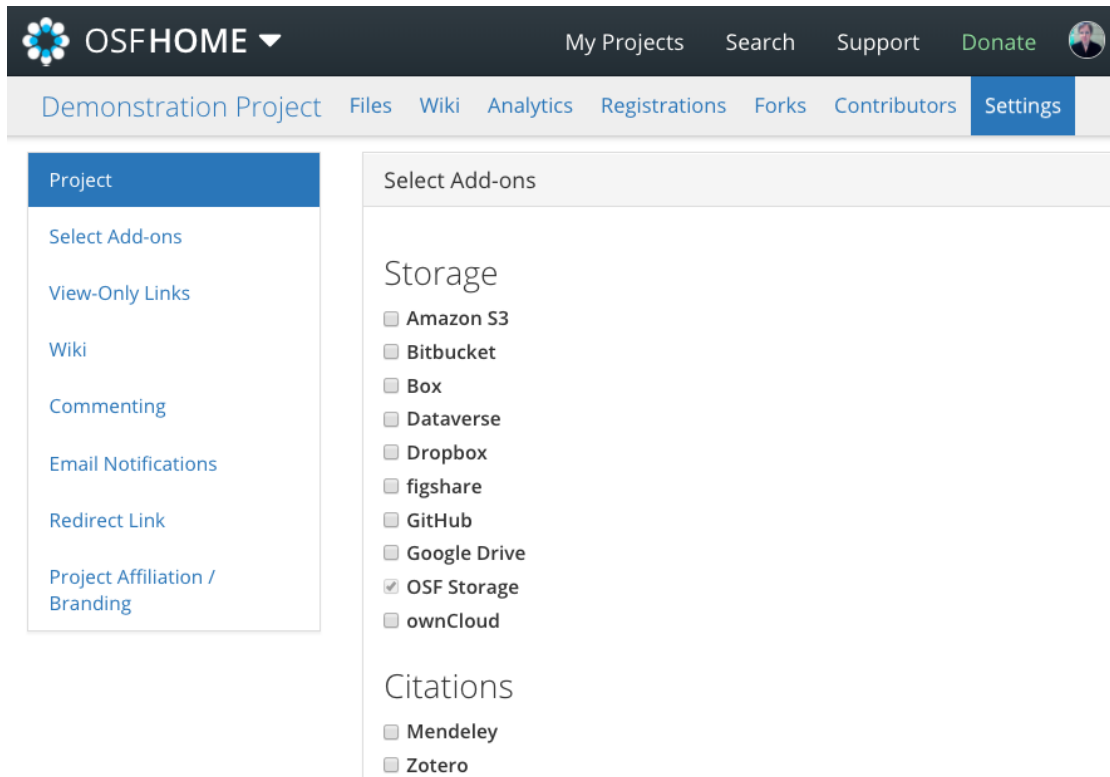
The image shows a screenshot of the Open Science Framework (OSF) website homepage. The header features the OSFHOME logo, a search bar, and links for Support, Donate, Sign Up, and Sign In. The main content area has a dark blue background with the text "Open Science Framework" and the tagline "A scholarly commons to connect the entire research cycle". Below this is a circular logo composed of white and blue dots. In the bottom right, there is a sign-up form with the heading "FREE AND OPEN SOURCE. START NOW." and fields for Full Name, Contact Email, Confirm Email, and Password (Must be 8 to 255 characters). A small inset window shows a project page with a play button overlay.

What it does

Connects various parts
of your workflow,
wherever they are

- Google Drive
- Dropbox
- Mendeley
- FigShare
- GitHub...

Share other non-project
files individually as
well (new feature)



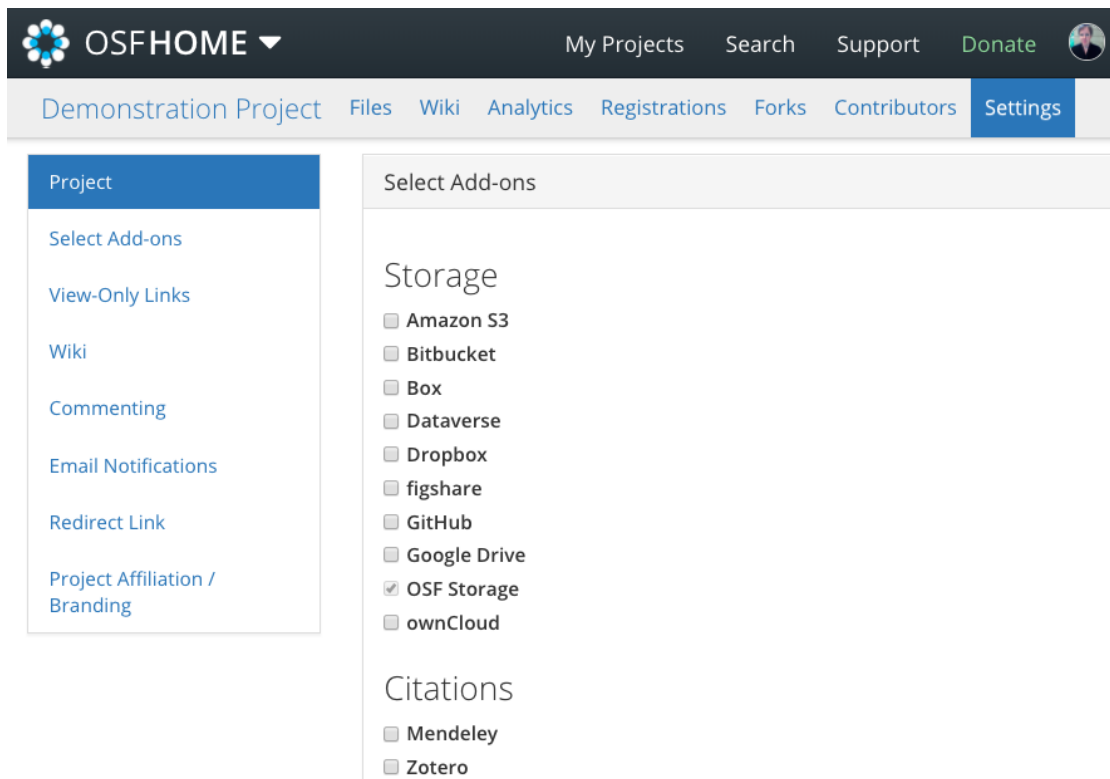
The screenshot displays the OSFHOME user interface. At the top, there is a dark navigation bar with the OSFHOME logo and a dropdown arrow on the left, and links for 'My Projects', 'Search', 'Support', and 'Donate' on the right, along with a user profile icon. Below this is a light gray navigation bar with links for 'Demonstration Project', 'Files', 'Wiki', 'Analytics', 'Registrations', 'Forks', 'Contributors', and 'Settings' (which is highlighted in blue). The main content area is divided into two columns. The left column, titled 'Project', contains a list of settings: 'Select Add-ons', 'View-Only Links', 'Wiki', 'Commenting', 'Email Notifications', 'Redirect Link', and 'Project Affiliation / Branding'. The right column, titled 'Select Add-ons', is further divided into 'Storage' and 'Citations' sections. The 'Storage' section lists various storage providers with checkboxes: Amazon S3, Bitbucket, Box, Dataverse, Dropbox, figshare, GitHub, Google Drive, OSF Storage (checked), and ownCloud. The 'Citations' section lists Mendeley and Zotero with checkboxes.

What it does

Supports versioning

Allows date-stamped registration of research projects

Provides an additional backup of research materials



The screenshot shows the OSFHOME interface for a project named "Demonstration Project". The top navigation bar includes "OSFHOM" with a logo, "My Projects", "Search", "Support", "Donate", and a user profile icon. Below this is a secondary navigation bar with "Demonstration Project" (selected), "Files", "Wiki", "Analytics", "Registrations", "Forks", "Contributors", and "Settings" (highlighted in blue). The main content area is divided into two columns. The left column, titled "Project", contains a list of settings: "Select Add-ons", "View-Only Links", "Wiki", "Commenting", "Email Notifications", "Redirect Link", and "Project Affiliation / Branding". The right column, titled "Select Add-ons", is further divided into "Storage" and "Citations" sections. The "Storage" section lists several options with checkboxes: Amazon S3, Bitbucket, Box, Dataverse, Dropbox, figshare, GitHub, Google Drive, OSF Storage (checked), and ownCloud. The "Citations" section lists Mendeley and Zotero, both with unchecked checkboxes.

What it does

Centralizes access to research information

Provides granular sharing of elements with collaborators

Provides access for others who can provide feedback at any stage of the research process

The screenshot shows the OSFHOME interface for a project named "Demonstration Project". The top navigation bar includes "OSFHOME", "My Projects", "Search", "Support", "Donate", and a user profile for "Amanda Izenstark". Below this, a secondary navigation bar highlights "Contributors" among other options like "Files", "Wiki", "Analytics", "Registrations", "Forks", and "Settings".

The main content area is titled "Contributors" and includes a "+ Add" button. Below the title, there is a "Filter by name" input field and a "Permissions" dropdown menu with options: "Administrator", "Read + Write", and "Read". There is also a "Bibliographic Contributor" dropdown with options: "Bibliographic" and "Non-Bibliographic".

The contributors are listed in two sections. The first section is titled "Contributors" and includes a "+ Add" button and the instruction "Drag and drop contributors to change listing order." It contains one contributor: Amanda Izenstark, with a profile picture, a "Permissions" dropdown set to "Administrator", a checked "Bibliographic Contributor" checkbox, and a "Remove" button.


The second section is titled "Admins on Parent Projects" and includes a "?" icon. It contains one contributor: Amanda Izenstark, with a profile picture, a "Permissions" dropdown set to "Read", an unchecked "Bibliographic Contributor" checkbox, and a "Remove" button.

At the bottom, there is a "View-only Links" section with a "+ Add" button and the instruction "Create a link to share this project so those who have the link can view—but not edit—the project."

Additional Related Project - OSF Preprints

Not just for science -
includes the Arts &
Humanities, Business,
Education, Law, and more.

* Once research is published,
encourage researchers to post
their final manuscripts your
institutional repository for
increased visibility!



The screenshot shows the OSF Preprints website homepage. At the top, there is a navigation bar with the OSF logo, the text "OSFPREPRINTS", and links for "Add a preprint", "Search", "Support", "Donate", "Sign Up", and "Sign in". Below the navigation bar is a large dark blue header area containing the OSF logo and the text "OSFPREPRINTS". A search bar is positioned in the center of the header, with the placeholder text "Search preprints..." and a "Search" button. Below the search bar, it states "2,123,274 searchable preprints as of November 14, 2017". Below this, there is a link "or" and a green button labeled "Add a preprint". A link "See an example" is also present. Below the header area, there is a section titled "Browse by subject" which contains a grid of ten blue buttons representing different subject areas: Architecture, Business, Engineering, Life Sciences, Physical Sciences and Mathematics, Arts and Humanities, Education, Law, Medicine and Health Sciences, and Social and Behavioral Sciences.

OSFPREPRINTS

Add a preprint Search Support Donate Sign Up Sign in

OSFPREPRINTS

Search preprints... Search

2,123,274 searchable preprints as of November 14, 2017

or

Add a preprint

See an example

Browse by subject

Architecture

Business

Engineering

Life Sciences

Physical Sciences and Mathematics

Arts and Humanities

Education

Law

Medicine and Health Sciences

Social and Behavioral Sciences

Closing thoughts

“As readers of scientific work, all we can do is be more skeptical of everything that is published.”

– Cristobal Young, Assistant Professor of Sociology, Stanford University, 2015

“I want to adopt a stance of humility and assume that there are errors and that’s why I need to be cautious in my conclusions.”

– Brian Nosek, Professor of Psychology, University of Virginia and co-founder and director of the Center for Open Science, 2016

Closing thoughts

Sharing research at various stages of the process for feedback and input from others can improve researchers' visibility, the actual research, and the final product.

(and in case you need additional talking points...)

A few things that would reduce stress around reproducibility/replicability in science

Jeff Leek 2017/11/21

I was listening to the Effort Report Episode on [The Messy Execution of Reproducible Research](#) where they were discussing the piece about [Amy Cuddy in the New York Times](#). I think both the article and the podcast did a good job of discussing the nuances of the importance of reproducibility and the challenges of the social interactions around this topic. After listening to the podcast I realized that I see a lot of posts about reproducibility/replicability, but many of them are focused on the technical side. So I started to think about compiling a list of more cultural things we can do to reduce the stress/pressure around the reproducibility crisis.

I'm sure others have pointed these out in other places but I am procrastinating writing something else so I'm writing these down while I'm thinking about them :).

- 1. We can define what we mean by “reproduce” and “replicate”** Different fields have different definitions of the words *reproduce* and *replicate*. If you are publishing a new study we now have an [R package](#) that you can use to create figures that show what changed and what was the same between the original study and your new work. Defining concretely what was the same and different will reduce some of the miscommunication about what a reproducibility/replicability study means.

<https://simplystatistics.org/2017/11/21/rr-sress/>

From “A few things...”

2. We can remember that replication is statistical, not deterministic

3. We can remember that there is a difference between exploratory and confirmatory research

6. We can be persistent and private as long as possible

7. We can make the realization that data is valuable but in science you don't own it

Thank you!

Andrée Rathemacher
Professor, Head of Acquisitions

andree@uri.edu

Amanda Izenstark
Professor, Reference & Instructional Design Librarian

amanda@uri.edu

Harrison Dekker
Associate Professor, Data Services Librarian

hdekker@uri.edu

