Northwestern Medicine®

Enhancing Reproducibility Through Good Computational and Data Practices

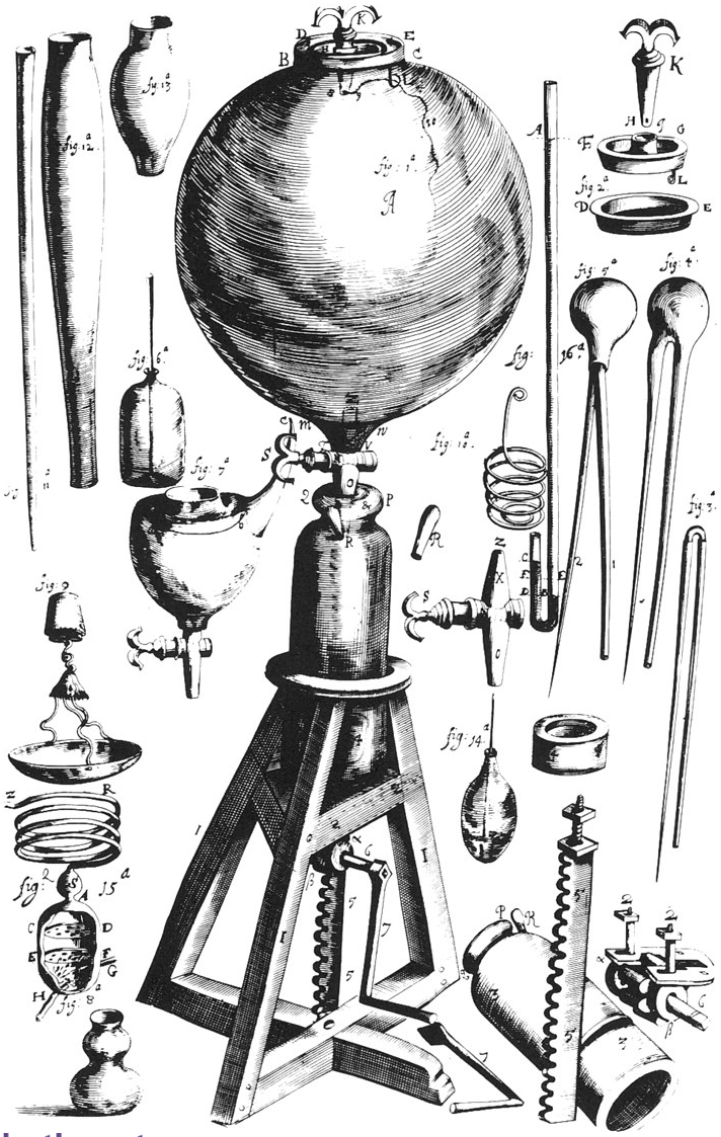
Matt Carson, PhD
Head, Digital Systems Department
Galter Health Sciences Library
matthew.carson@northwestern.edu

Questions to Answer

- What is “reproducibility”?
- Why should we care about reproducibility?
- What are the incentives (besides just doing the right thing)?
- What can we do to enhance reproducibility in our own work?
- What resources are available here at NU to support reproducible research?

What is “reproducibility”?

Robert Boyle's Vacuum Project (1660's)



- A complicated and expensive piece of machinery
- Reproducing his results was difficult
- Boyle made a case that empirical findings must be verified by independent replication
- Wanted to provide sufficient details on **procedure, protocols, equipment, and observations** so…

  "that the person I addressed them to might, without mistake, and with as little trouble as possible, be able to repeat such unusual experiments".

https://en.wikipedia.org/wiki/Reproducibility#/media/File:Boyle_air_pump.jpg

Reproducible Science Today

What is it? What's required to document it?

"***reproducibility*** refers to the ability of a researcher to duplicate the results of a prior study using the same materials as were used by the original investigator. That is, a second researcher might use the same raw data to build the same analysis files and implement the same statistical analysis in an attempt to yield the same results.... ***Reproducibility is a minimum necessary condition for a finding to be believable and informative***."

Social, Behavioral, and Economic Sciences Perspectives on Robust and Reliable Science (National Science Foundation, Arlington, VA, 2015).

REQUIRES SHARING **DATA, METADATA, ANALYTICAL CODE, AND SOFTWARE**
...AT A MINIMUM

Science Translational Medicine 01 Jun 2016: Vol. 8, Issue 341, pp. 341ps12
DOI: 10.1126/scitranslmed.aaf5027

Northwestern Medicine®

Types of Reproducibility

- Empirical
  - Boyle's case for communicating and sharing experimental info
- Statistical
  - Detailed information provided about statistical tests used, model parameters, threshold values, etc. (i.e., to prevent p-hacking)
- **Computational**
  - **Requires access to code, data, and implementation details**

A problem with any of these three types can prevent the establishment of scientific fact

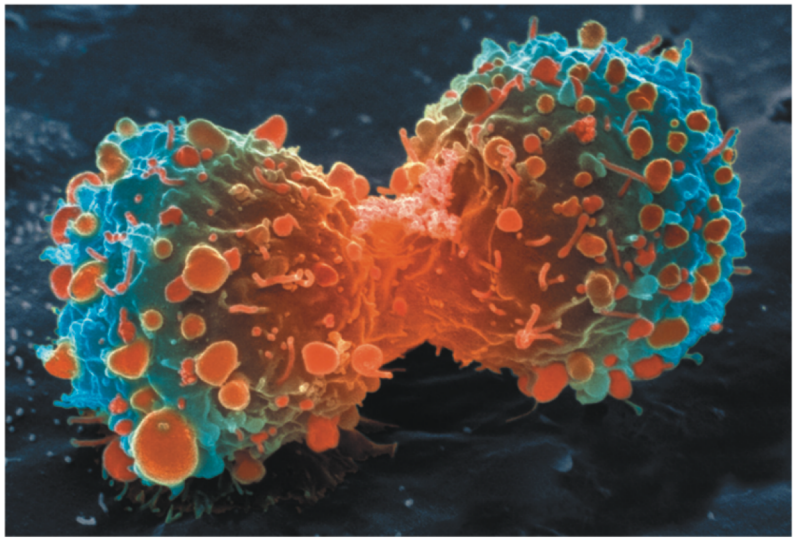
Victoria Stodden, 2014

Barriers to Reproducibility

- For many journal articles, associated data is not available
- Difficulty of accessing unpublished data
- Few avenues for publishing negative results
- Failure of funding agencies to establish or enforce data access policies

Why should we care about reproducibility?

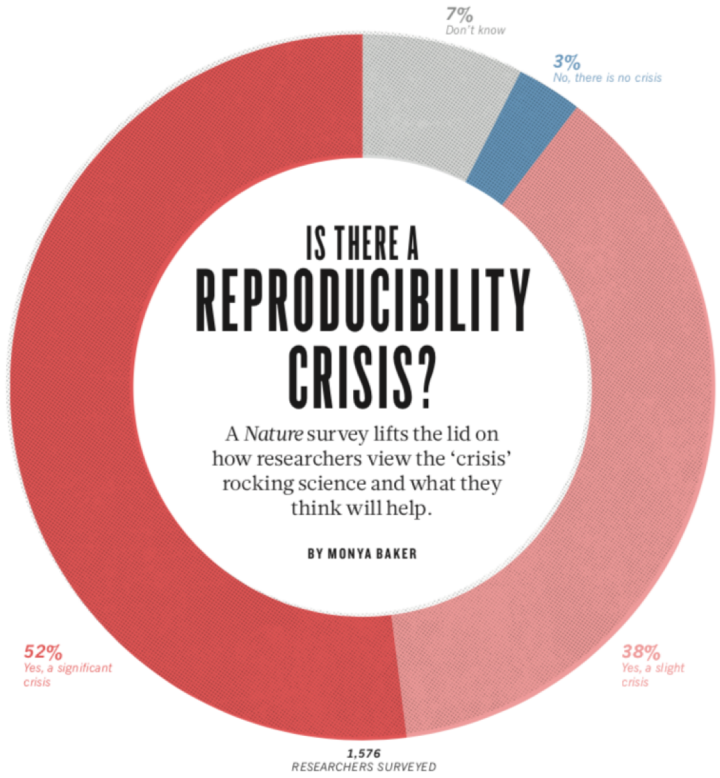
Lack of Reproducibility Hinders Patient Benefits



Many landmark findings in preclinical oncology research are not reproducible, in part because of inadequate cell lines and animal models.

Raise standards for preclinical cancer research

C. Glenn Begley and **Lee M. Ellis** propose how methods, publications and incentives must change if patients are to benefit.



1. Begley & Ellis. *Nature,*2012
2. Baker, *Nature*, 2018

Northwestern Medicine®

In many cases, published articles do not provide enough detail for the reader to recreate the experiment

Table 4: Computational Reproducibility Evaluation (n=55)

Straightforward to reproduce with minimal effort	0%
Minor difficulty in reproducing	0%
Reproducible after some tweaking	9.1%
Could reproduce with fairly substantial skill and knowledge	16.4%
Reproducible with substantial intellectual effort	12.7%
Reproducible with substantial tedious effort	3.6%
Difficult to reproduce because of unavoidable inherent complexity	3.6%
Nearly impossible to reproduce	3.6%
Impossible to reproduce	50.9%

Stodden, Krafczyk, and Bhaskar. "Enabling the Verification of Computational Results", 2018.

What are the incentives (besides just doing the right thing)?

Incentives for Reproducibility

- To provide evidence that your results are correct
- To allow other researchers to reuse your methods and/or results
- Preservation and sharing increasingly required by…
  - Federal funding agencies
  - Journals
- Journals hiring "reproducibility editors"
- More scholars calling for greater standards for reproducibility
- The researcher you are helping most is **your future self**

National Institutes of Health
Turning Discovery Into Health

Search NIH

NIH Employee Intranet | Staff Directory | En Español

Health Information | Grants & Funding | News & Events | Research & Training | Institutes at NIH | About NIH

Home » Research & Training » Rigor and Reproducibility

RIGOR AND REPRODUCIBILITY

Principles and Guidelines for Reporting Preclinical Research (2014)

- Rigorous statistical analysis
- Transparency in reporting
- Data and material sharing
- Consideration of refutations
- Consider establishing best practice guidelines for images and reagents

Updated Application Instructions to Enhance Rigor and Reproducibility (2016)

- Scientific Premise of Proposed Research
- Rigorous Experimental Design
- Consideration of Sex and Other Relevant Biological Variables
- Authentication of Key Biological and/or Chemical Resources

From: https://www.nih.gov/research-training/rigor-reproducibility

FAIR Principles for Discoverability

FAIR Principles put specific emphasis on enhancing the ability of machines to automatically find and use the data

https://libereurope.eu

Data and supplementary materials have sufficiently rich metadata and a unique and persistent identifier.

FINDABLE

Metadata and data are understandable to humans and machines. Data is deposited in a trusted repository.

ACCESSIBLE

Metadata use a formal, accessible, shared, and broadly applicable language for knowledge representation.

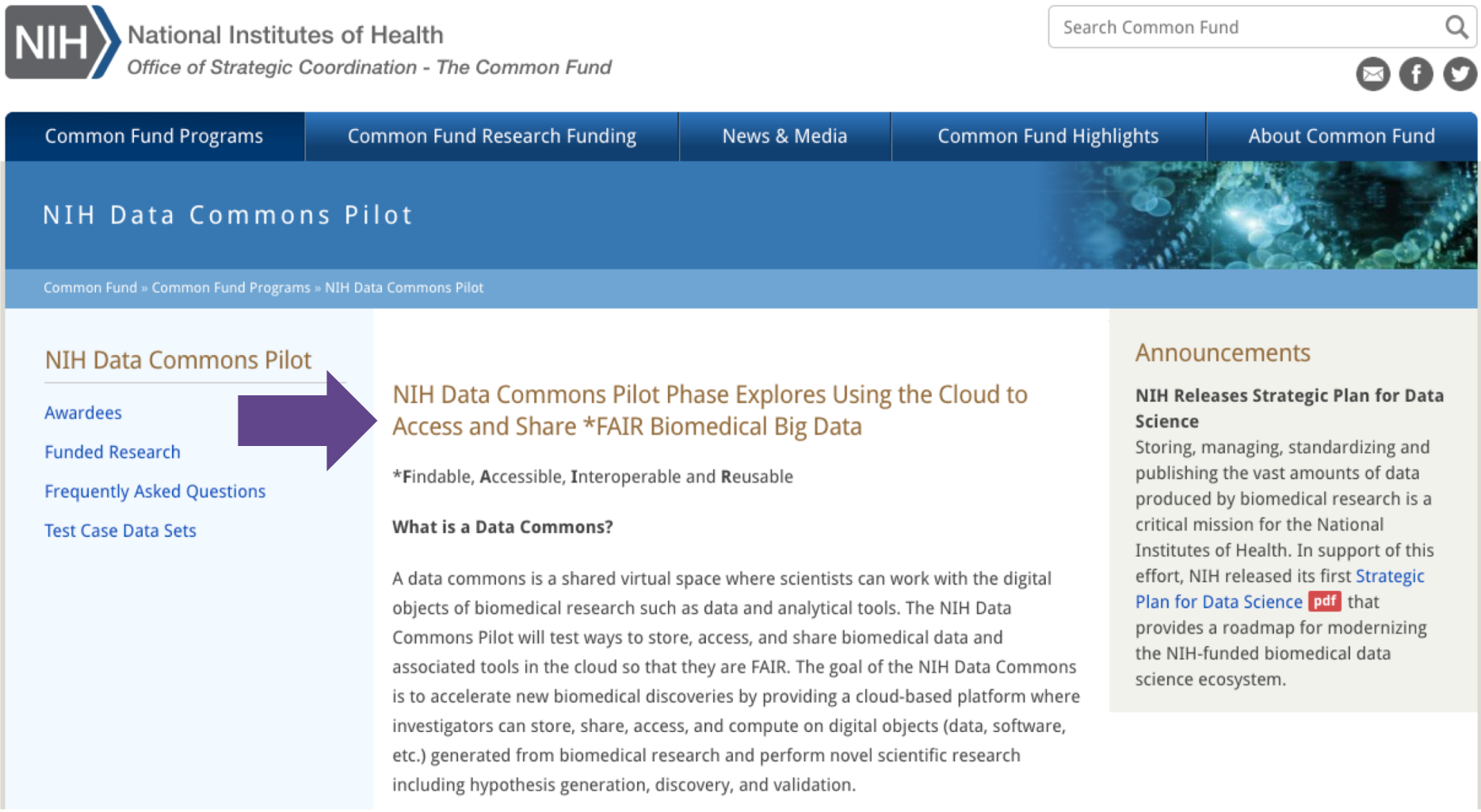
INTEROPERABLE

Data and collections have a clear usage licenses and provide accurate information on provenance.

REUSABLE

Wilkinson et al., "The FAIR Guiding Principles for scientific data management and stewardship", 2016

NIH Data Commons Supporting FAIR Principles



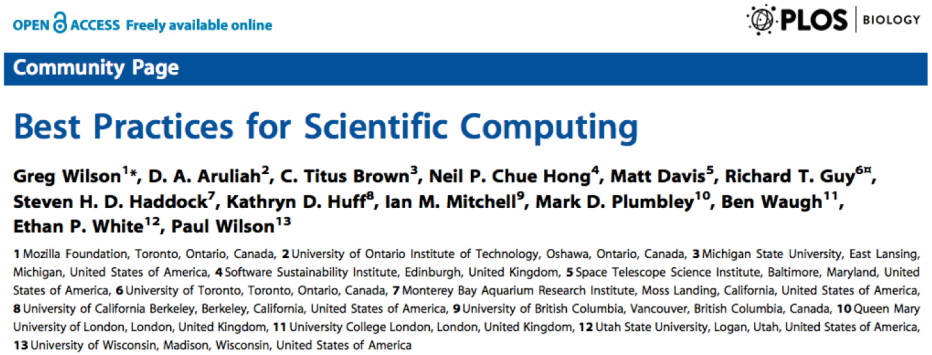
https://commonfund.nih.gov/commons

Northwestern Medicine®

What can we do to enhance (computational) reproducibility in our own work?

Scientific Computing and Data Management

Best Practices



OPEN ACCESS Freely available online

PLOS | BIOLOGY

Community Page

Best Practices for Scientific Computing

Greg Wilson[1]*, D. A. Aruliah[2], C. Titus Brown[3], Neil P. Chue Hong[4], Matt Davis[5], Richard T. Guy[6¤], Steven H. D. Haddock[7], Kathryn D. Huff[8], Ian M. Mitchell[9], Mark D. Plumbley[10], Ben Waugh[11], Ethan P. White[12], Paul Wilson[13]

1 Mozilla Foundation, Toronto, Ontario, Canada, **2** University of Ontario Institute of Technology, Oshawa, Ontario, Canada, **3** Michigan State University, East Lansing, Michigan, United States of America, **4** Software Sustainability Institute, Edinburgh, United Kingdom, **5** Space Telescope Science Institute, Baltimore, Maryland, United States of America, **6** University of Toronto, Toronto, Ontario, Canada, **7** Monterey Bay Aquarium Research Institute, Moss Landing, California, United States of America, **8** University of California Berkeley, Berkeley, California, United States of America, **9** University of British Columbia, Vancouver, British Columbia, Canada, **10** Queen Mary University of London, London, United Kingdom, **11** University College London, London, United Kingdom, **12** Utah State University, Logan, Utah, United States of America, **13** University of Wisconsin, Madison, Wisconsin, United States of America

PERSPECTIVE

Good enough practices in scientific computing

Greg Wilson[1☯]*, Jennifer Bryan[2☯], Karen Cranston[3☯], Justin Kitzes[4☯], Lex Nederbragt[5☯], Tracy K. Teal[6☯]

1 Software Carpentry Foundation, Austin, Texas, United States of America, **2** RStudio and Department of Statistics, University of British Columbia, Vancouver, British Columbia, Canada, **3** Department of Biology, Duke University, Durham, North Carolina, United States of America, **4** Energy and Resources Group, University of California, Berkeley, Berkeley, California, United States of America, **5** Centre for Ecological and Evolutionary Synthesis, University of Oslo, Oslo, Norway, **6** Data Carpentry, Davis, California, United States of America

☯ These authors contributed equally to this work.
* gvwilson@software-carpentry.org

OPEN ACCESS Freely available online

PLOS | COMPUTATIONAL BIOLOGY

Editorial

Ten Simple Rules for Reproducible Computational Research

Geir Kjetil Sandve[1,2]*, Anton Nekrutenko[3], James Taylor[4], Eivind Hovig[1,5,6]

1 Department of Informatics, University of Oslo, Blindern, Oslo, Norway, **2** Centre for Cancer Biomedicine, University of Oslo, Blindern, Oslo, Norway, **3** Department of Biochemistry and Molecular Biology and The Huck Institutes for the Life Sciences, Penn State University, University Park, Pennsylvania, United States of America, **4** Department of Biology and Department of Mathematics and Computer Science, Emory University, Atlanta, Georgia, United States of America, **5** Department of Tumor Biology, Institute for Cancer Research, The Norwegian Radium Hospital, Oslo University Hospital, Montebello, Oslo, Norway, **6** Institute for Medical Informatics, The Norwegian Radium Hospital, Oslo University Hospital, Montebello, Oslo, Norway

EDITORIAL

Ten Simple Rules for Digital Data Storage

Edmund M. Hart[1]*, Pauline Barmby[2], David LeBauer[3], François Michonneau[4,5], Sarah Mount[6], Patrick Mulrooney[7], Timothée Poisot[8], Kara H. Woo[9], Naupaka B. Zimmerman[10], Jeffrey W. Hollister[11]

1 University of Vermont, Department of Biology, Burlington, Vermont, United States of America, **2** University of Western Ontario, Department of Physics and Astronomy, London, Canada, **3** University of Illinois at Urbana-Champaign, National Center for Supercomputing Applications and Institute for Genomic Biology, Urbana, Illinois, United States of America, **4** University of Florida, iDigBio, Florida Museum of Natural History, Gainesville, Florida, United States of America, **5** University of Florida, Whitney Laboratory for Marine Bioscience, Gainesville, Florida, United States of America, **6** King's College London, Department of Informatics, London, United Kingdom, **7** University of California at San Diego, San Diego Supercomputer Center, San Diego, California, United States of America, **8** Université de Montréal, Département de Sciences Biologiques, Montreal, Canada, **9** Washington State University, Center for Environmental Research, Education, and Outreach, Pullman, Washington, United States of America, **10** University of Arizona, School of Plant Sciences, Tucson, Arizona, United States of America, **11** US Environmental Protection Agency, Atlantic Ecology Division, Narragansett, Rhode Island, United States of America

* edmund.m.hart@gmail.com

Northwestern Medicine®

Good Enough Practices in Scientific Computing

Data Management

- Save raw data
- Use multiple backup locations (e.g., Box, other off-site storage)
- Create the data you wish to see
  - Convert to open formats (.csv, JSON, XML, etc.)
  - Use clear variable names ('*name1*' and '*name2*' to '*personal_ name*' and '*family_name*')
  - Store useful metadata as part of the file name (e.g., *2016-05-alaska-b.csv*)
  - (*Create a data dictionary*)
- Create analysis-friendly or "tidy" data (next slide)
- Record all of the steps used to process your data
  - write scripts for *every* stage of data processing
  - Allows for much faster reproduction later **(direct benefits future you)**
- Use unique identifiers for *every* record (in case you have multiple tables)
- Submit data to a DOI-issuing repository so others can cite and access it

Tidy Data Example

Make each column a variable, make each row an observation

https://www.tidyverse.org/

site	1999	2000
Whitehorse	745	2666
Yellowknife	37737	80488
Inuvik	212258	213766

"untidy"

site	year	cases
Whitehorse	1999	745
Whitehorse	2000	2666
Yellowknife	1999	37737
Yellowknife	2000	80488
Inuvik	1999	212258
Inuvik	2000	213766

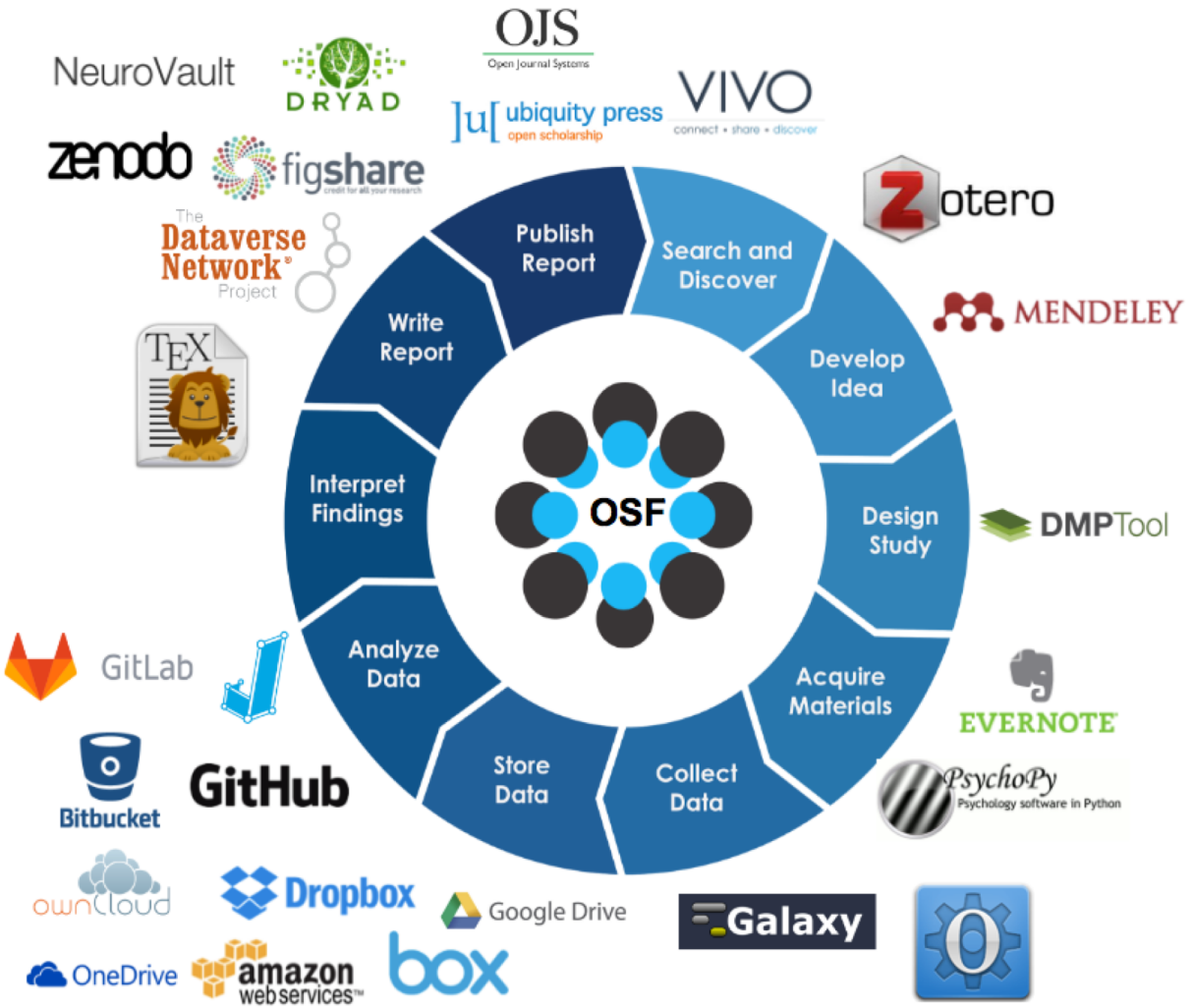
tidy

1. Wilson, et al. "Good Enough Practices in Scientific Computing", 2017. https://doi.org/10.1371/journal.pcbi.1005510
2. Hadley Wickham, "Tidy Data", 2014. http://dx.doi.org/10.18637/jss.v059.i10

Tools for Supporting Reproducible Research

- **Project Organization and Collaboration**
  - Open Science Framework (as of July 16 OSF has surpassed 100,000 users)
  - Lab notebooks (Benchling, https://benchling.com)
  - Slack (https://slack.com)
- **Data Collection**
  - RedCap (for clinical research)
- **Data analysis and documentation in computational notebooks**
  - R/Rstudio/Knitr
  - Python/Jupyter Notebooks
- **Manuscript Preparation**
  - R Markdown
  - StatTag
- **Backup and Storage** (Box, off-site storage)
- **Preservation**
  - Institutional or subject repository (DigitalHub)
  - NIH Data Sharing Repos

Open Science Framework



Lab Example

Public | 5

Contributors: Ian Sullivan, Courtney K. Soderberg, Jennifer Freeman Smith, Brandon Thorpe

Affiliated institutions: Center For Open Science

Date created: 2017-06-06 11:00 AM | Last Updated: 2018-03-22 02:41 PM

Category: Project

Description: This project demonstrates one way the OSF can be used to create a space for a lab to share materials and research.

License: CC0 1.0 Universal

Wiki

This is an example project showing how the OSF might be used by a lab to create a shared lab space, share lab standards and resources, and collate the work that is being done by individuals/groups within the lab.

Getting started

Welcome to the lab! This wiki contains the steps to take at key moments in the life of a lab like when someone joins the lab, when you start a new experiment or when the p...

Read More

Files

Filter

Name	Modified
Lab Example	
– OSF Storage	
– New Experiment Template	
– OSF Storage	
+ Data	
+ Analysis Scripts	
+ Manuscript	
+ Experiment Protocol	
– Lab Documents	
– OSF Storage	
figure-copyTemplate.png	2017-06-06 07:20 PM
– Lab Meetings	
– OSF Storage	
+ 2016-Fall	

Citation

osf.io/n7263

Components

- New Experiment Template
  Sullivan, Soderberg, Smith & 1 more
  Start all new experiments by forking this section.
- Lab Documents
  Sullivan, Soderberg, Smith & 1 more
  Internal documentation and lab reference material should go here.
- Lab Meetings
  Sullivan, Soderberg, Smith & 1 more
  Discussion notes, articles, and other internal musings.
- Research
  Sullivan, Soderberg, Smith & 1 more
  Link all your experiment projects here.
- OSF video for Cambridge ELN pilot
  Soderberg

Tags

demonstration | example project | OSF Example | tutorial

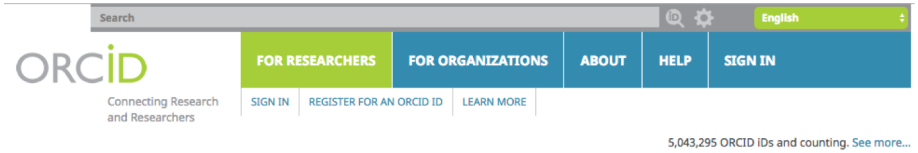
Recent Activity

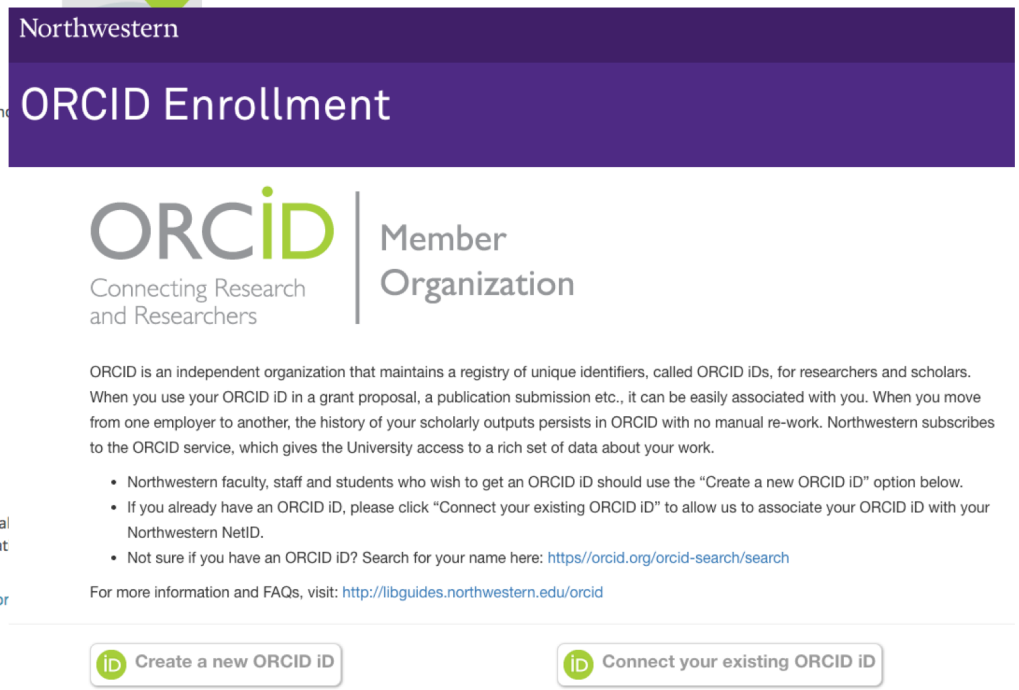
Ian Sullivan created a link to Second Experiment
2018-04-18 10:42 AM

Ian Sullivan edited description of Lab Documents
2018-04-18 10:18 AM

Open Researcher and Contributor iD (ORCiD)

https://orcid.it.northwestern.edu/





https://galter.northwestern.edu/guides-and-tutorials/orcid?category=26

What resources are available here at NU to support reproducible research?

Research Data Management Support

Northwestern Resources

- **Across campus**
  - DMPTool (for PIs)
  - Arch (NUL Repository)
- **Galter Library**
  - Digital Hub (Repository for Feinberg)
  - Data Community Engagement
  - Next-generation Integrated Repository and Data Index

Data Management Plan

- Increasingly required by funding agencies

- A document that describes *what you are going to do with your data during your research project and after you complete it*

- What does it include?
  1. Types of data, how it will be collected, how it will be processed
  2. What metadata is needed to make the data meaningful and how will you capture it?
  3. Policies for access and sharing
  4. Policies for reuse and redistribution
  5. Plans for archiving and preservation

Learn | Sign in | English (US)

DMPTool
Build your Data Management Plan

DMPTool by the Numbers

30,753
Users

27,281
Plans
More

236
Participating Institutions
More

Top 5 Templates

Digital Curation Centre
NSF-BIO: Biological Sciences
NSF-SBE: Social, Behavioral, Economic Sciences
USDA-NIFA: National Institute of Food and Agriculture
NIH-GEN: Generic

More

DMPTool News

Scoping Machine-Actionable DMPs

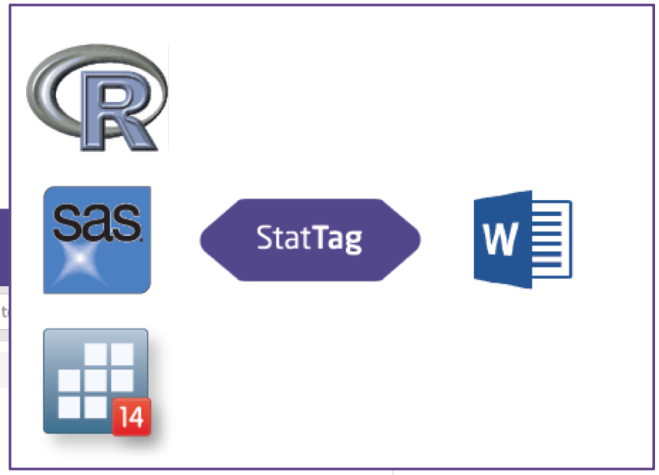
Go to the blog
RSS

About | Terms of use | Privacy statement | Accessibility | Github | Contact us

DMPTool

DigitalHub: StatTag

https://digitalhub.northwestern.edu



DigitalHub

Home | Catalog | About | News | Collections | Help | Contact | Enter search t

Back to search results

Download the file

Actions

Download | Analytics | Citations

Export to: EndNote

2

3

Collections

This file is not currently in any collections.

StatTag

Open Access (recommended)

Related URL refers to the software: www.stattag.org

Descriptions

Resource type(s): Software or Program Code

Keyword: StatTag, Reproducible Research, Stata, SAS, R, Statistical Code, Plug-in

Rights: The MIT License (MIT)

Creator: Welty, Leah J; Rasmussen, Luke Vincent; Baldridge, Abigail Shubat; Whitley, Eric

Abstract: StatTag is a free plug-in for conducting reproducible research and creating dynamic documents using Microsoft Word with the Stata and SAS statistical software (future versions will work with R). StatTag allows users to embed statistical output (estimates, tables, and figures) within Word and provides an interface to edit statistical code directly from Word. Statistical output can be individually or collectively updated from Word in one-click with a behind-the-scenes call to the statistical program. With StatTag, modification of a dataset or analysis no longer entails transcribing or re-copying results in to a manuscript or table.

Related URL: www.stattag.org

Digital Publisher: Galter Health Sciences Library, Feinberg School of Medicine, Northwestern University

Date Created: 2016-07-01

Language: English

Subject: MESH: Statistics as Topic

Subject: LCSH: Plug-ins (Computer programs); Statistics--Data processing; R (Computer program language)

Subject: Name: Stata; SAS (Computer file)

Location: Chicago, Illinois, United States

DOI: 10.18131/G36K76

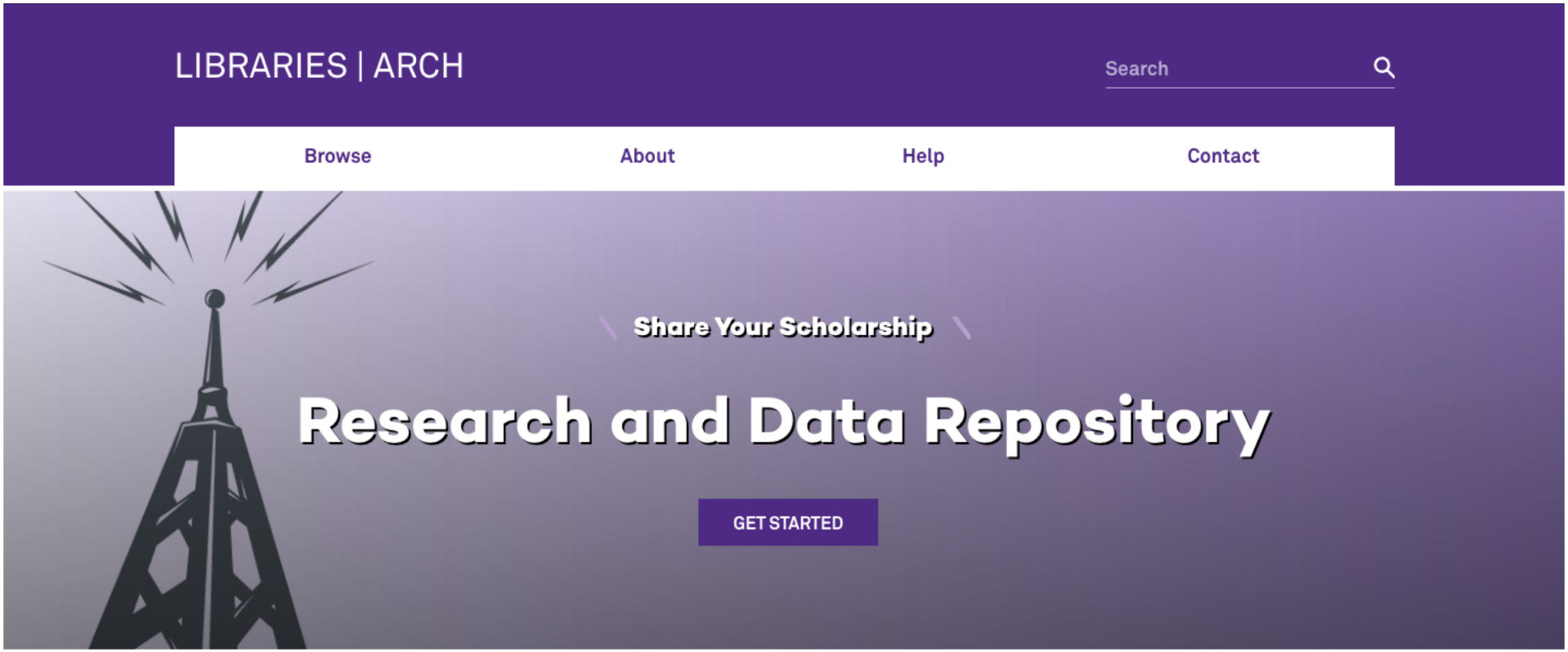
ARK: ark:/c8131/g36k76

Northwestern Medicine®

http://sites.northwestern.edu/stattag/

Arch Research and Data Repository

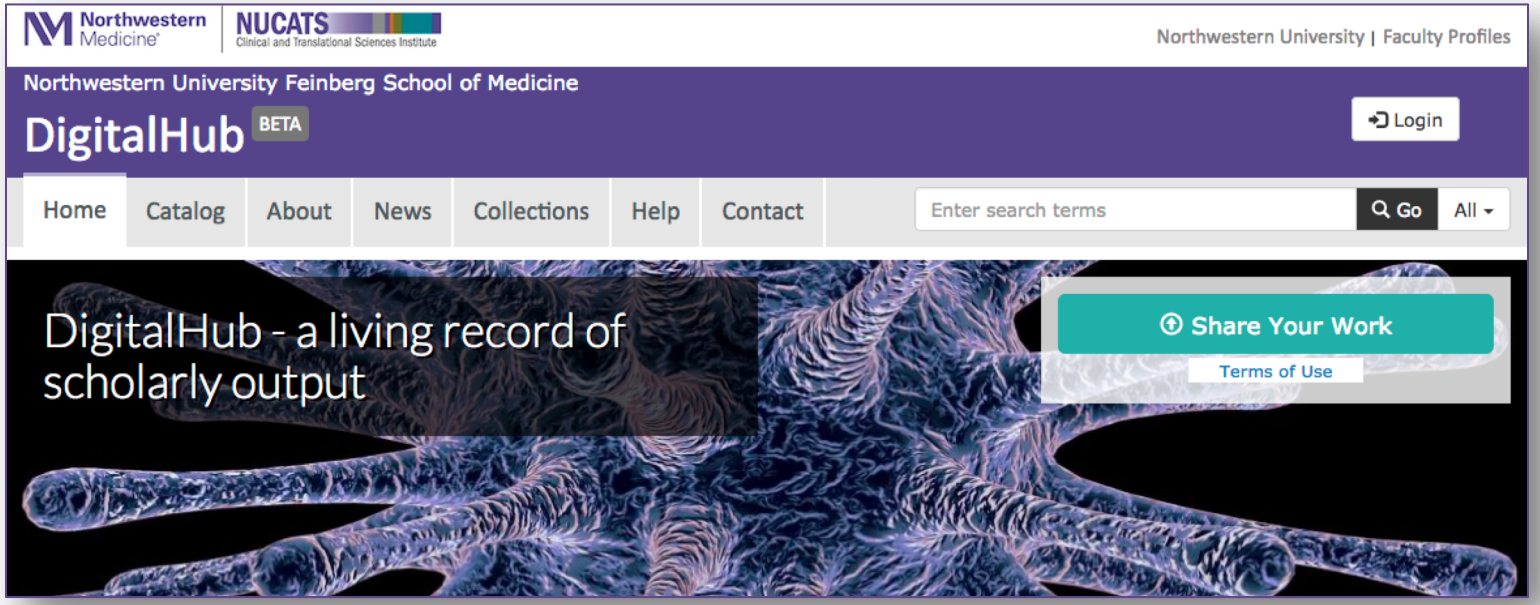
Northwestern University Library



arch.library.northwestern.edu

Galter Library Resources

The Repository -- Research Preservation, Dissemination, and Impact: DigitalHub *https://digitalhub.northwestern.edu*



- Digital Object Identifiers supplied for a wide variety of deposited objects provide secure and long term preservation, easy citation and searchability, and persistent access.
- Metrics for deposited items, management of licensing, enhanced discoverability, and support for large datasets available

Galter Library: Data Community Engagement

The Galter Library actively promotes reproducibility and open science best practices by organizing, hosting, and/or sponsoring special events featuring cutting-edge organizations and software tools in the field.

- **Workshops**
  - *R: Refresher and Visualization with ggplot2 Workshop* ***(NUIT Res Comp)*** [July 2018]
  - *Integrating Reproducible Best Practices into Biomedical and Clinical Research: A Hands-on Workshop for Researchers* ***(Code Ocean)*** [April 2018]
  - *Computational Skills for Informatics Series* [Winter 2018]
  - *Practical Steps for Increasing Openness and Reproducibility: An Introduction to the* ***Open Science Framework (OSF)*** [November 2017]
- **Hackathons**
  - *Post-ISMB 2018 Chicago Bioinformatics Hackathon* ***(NCBI)*** [July 2018]

A next-gen integrated repository infrastructure

Guiding Principles

- ✓ **Distribution of control**

  Distributed control, or governance, of scholarly resources are more sustainable and at less risk to buy-out or failure.

- ✓ **Inclusiveness and diversity**

  Different institutions and regions have unique and particular needs and contexts (e.g. diverse language, policies, and priorities). A distributed network of repositories will aim to reflect and be responsive to the different needs and contexts of different regions, disciplines and countries.

- ✓ **Public good**

  The technologies, architectures and protocols adopted in the context of the global network for repositories will be available to everyone, using global standards when they are available.

- ✓ **Intelligent openness and accessibility**

  Scholarly resources will be made openly available and in accessible formats, whenever possible, in order increase their value and maximize their re-use for the benefit for scholarship and society.

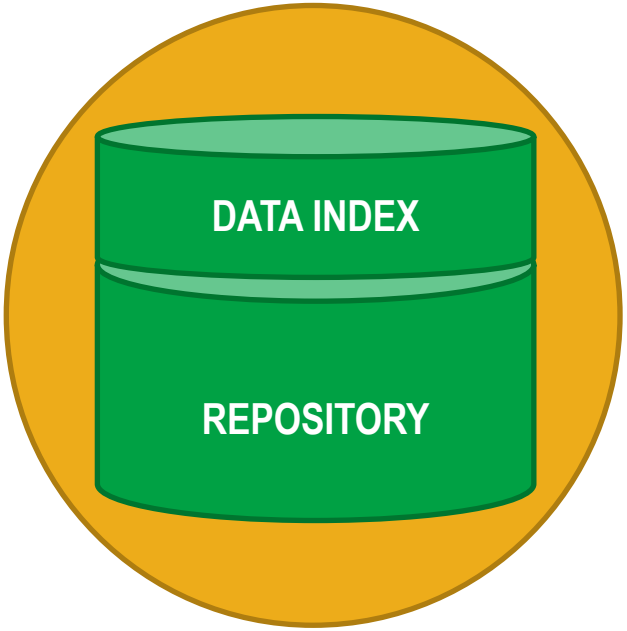
- ✓ **Sustainability**

  Institutions and research organizations will be major participants in the global network, contributing to the long term sustainability of resources.

- ✓ **Interoperability**

  Repositories will adopt common behaviors, functionalities and standards ensuring interoperability across institutions and enabling them to engage in a common way with external service providers

Integrated repository & data index

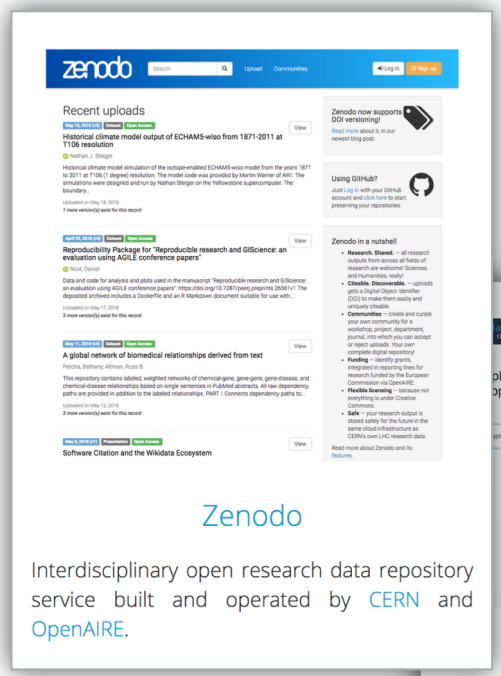


VISION:

*A foundation for a **distributed, globally networked infrastructure** on top of which layers of **value added services** can be deployed, making it more **research-centric, open to and supportive of innovation**, while also c**ollectively managed by the scholarly community.***

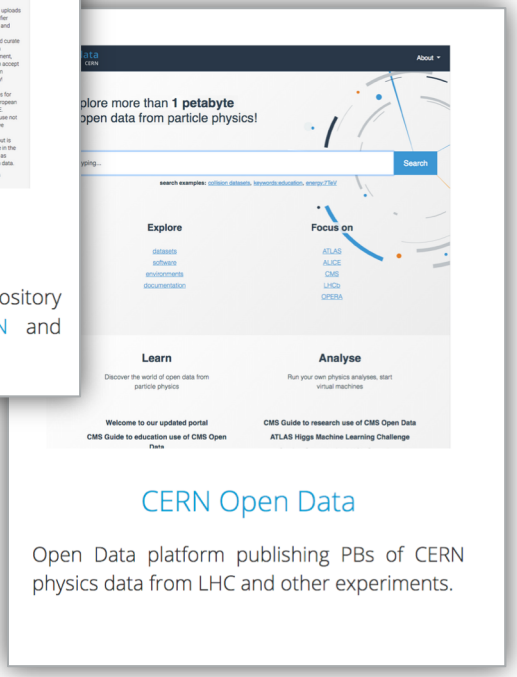
Make the resource, rather than the repository, the focus of services and infrastructure.

Leveraging Invenio 3.0 as a strong foundation



Zenodo

Interdisciplinary open research data repository service built and operated by CERN and OpenAIRE.



CERN Open Data

Open Data platform publishing PBs of CERN physics data from LHC and other experiments.

Safe: Invenio has been created with security and long-term preservation in mind.

Scalable: Invenio is fast. Designed to manage 100+ million records and petabytes of files. All your research data can now be archived independently of the size.

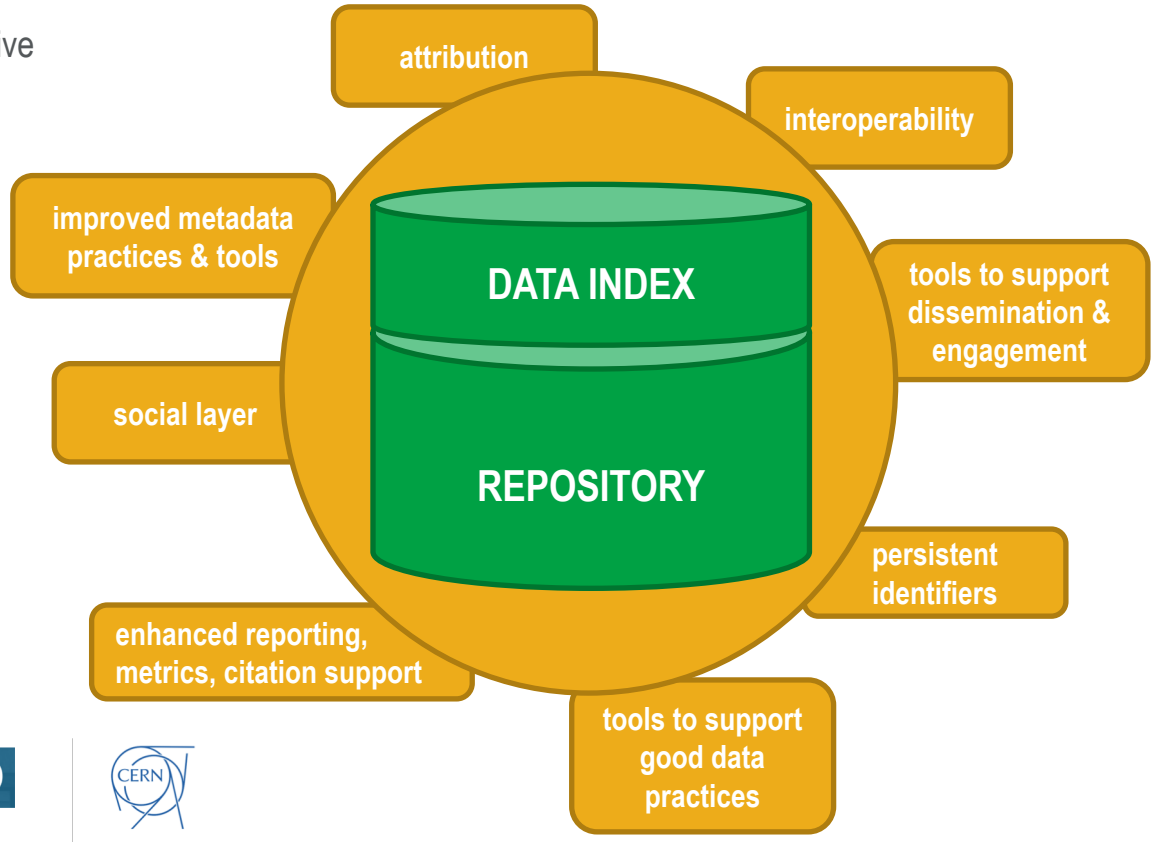
RESTful: Only a modern framework can create modern digital repositories. Invenio was born for the web, is JSON-native and provides RESTful APIs out of the box that will allow to build apps on top of it.

Open: Invenio is 100% open source licensed under MIT license. Invenio loves open standards for open science.

Contribute to the NGR collaboration + enhanced features

NGR: collaborative work to facilitate the development of new services on top of the collective network.

1. Exposing Identifiers
2. Declaring Licenses at the Resource Level
3. Discovery Through Navigation
4. Interacting with Resources (Annotation, Commentary, and Review)
5. Resource Transfer
6. Batch Discovery
7. Collecting and Exposing Activities
8. Identification of Users
9. Authentication of Users
10. Exposing Standardized Usage Metrics
11. Preserving Resources



COAR Confederation of Open Access Repositories

NEXT GENERATION REPOSITORIES

INVENIO

NII 国立情報学研究所 National Institute of Informatics

CERN technology

https://github.com/inveniosoftware/invenio

Resources

- Challenges in Irreproducible Research. Specials and Supplements Archive, Nature. http://www.nature.com/news/reproducibility-1.17552
- NIH Rigor & Reproducibility. Available: https://www.nih.gov/research-training/rigor-reproducibility
- NIH Data Commons https://commonfund.nih.gov/bd2k/commons
- Council of Councils meeting held on 26 May 2017 https://dpcpsi.nih.gov/council/05262017agenda incl. Update on Plans of the NIH Data Commons (Vivien Bonazzi, Ph.D., Program Director, OSC, DPCPSI) https://dpcpsi.nih.gov/sites/default/files/CoCMay2017_320PM_UpdateOnNIHDataCommons.pdf
- NIH Data Sharing Repositories: https://www.nlm.nih.gov/NIHbmic/nih_data_sharing_repositories.html
- DigitalHub: https://digitalhub.northwestern.edu
- StatTag: http://sites.northwestern.edu/stattag/

Data Policies and Procedures

Funding Agencies and Journals

- NIH Data Sharing Policies: https://www.nlm.nih.gov/NIHbmic/nih_data_sharing_policies.html
- The Scholarly Publishing and Academic Resources Coalition (SPARC*) Data Sharing Policies Comparison Tool: http://datasharing.sparcopen.org/
- PLOS Data Sharing Policy: http://journals.plos.org/plosone/s/materials-and-software-sharing
- Data De-Identification Policies (HHS): https://www.hhs.gov/hipaa/for-professionals/privacy/special-topics/de-identification/index.html

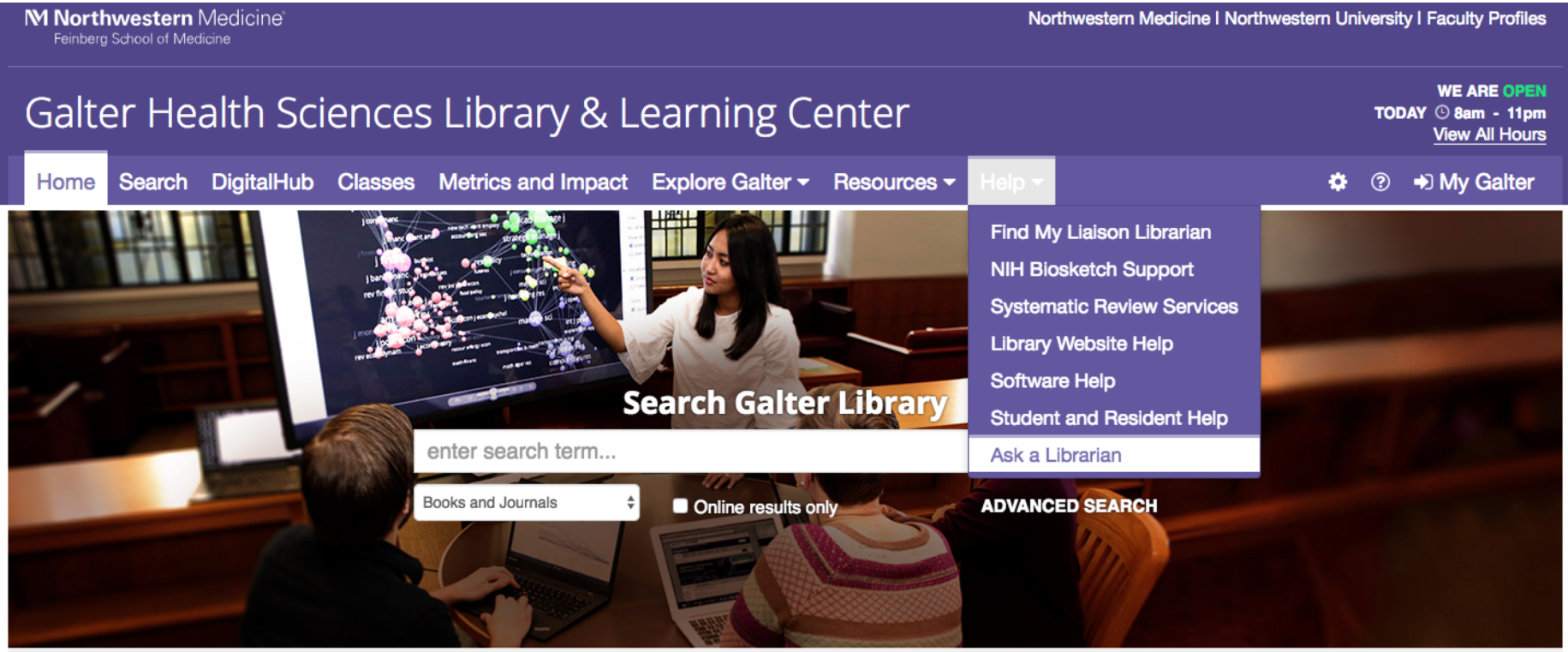
NU Services Supporting Best Practices

- **Galter Health Sciences Library & Learning Center**: https://galter.northwestern.edu/
- **Northwestern University Libraries**: https://www.library.northwestern.edu/
- **NUIT Research Computing**: http://www.it.northwestern.edu/research/index.html
- **Biostatistics Collaboration Core (BCC)**: http://www.feinberg.northwestern.edu/sites/bcc/
- **REDCap support**: https://nucats.northwestern.edu/resources-services/data-informatics-services/software-tools/redcap

Data Policies and Procedures (NU)

- **Institutional Review Board (IRB)**: https://irb.northwestern.edu/
  - Data Review Protocol: https://irb.northwestern.edu/templates-forms/templates-forms-sops
- **Office of Sponsored Research (OSR)**: https://osr.northwestern.edu/
  - Data Use Agreement (DUA): https://osr.northwestern.edu/agreements/dua
  - Data Retention Policy: https://osr.northwestern.edu/sites/osr/files/research_data.pdf
- **FSMIT**: http://www.feinberg.northwestern.edu/it/
  - Information Security: http://www.feinberg.northwestern.edu/it/policies/information-security/index.html
  - Security Policies: http://www.feinberg.northwestern.edu/it/policies/index.html
  - Guidelines for files storage: http://www.feinberg.northwestern.edu/it/policies/file-storage.html
  - Data Security Plan for information used in clinical research: http://www.feinberg.northwestern.edu/it/policies/information-security/data-security-plans.html
- **Data Management Plan (DMP) Tool**:
  - https://dmptool.org (Links to an external site.)Links to an external site.

Galter Health Sciences Library & Learning Center



You can also email me with any questions:

matthew.carson@northwestern.edu

Northwestern Medicine

Contributors to this Presentation

Galter Health Sciences Library & Learning Center

Matt Carson

Head, Digital Systems

Senior Research Data Scientist

matthew.carson@northwestern.edu

Pamela Shaw

Biosciences &

Bioinformatics Librarian

p-shaw2@northwestern.edu

Kristi Holmes

Director

Associate Professor of Preventive Medicine

(Health and Biomedical Informatics) and

Medical Education

Northwestern Medicine®

Thank you!