

# Research Data Management

Ellie Ransom

Research Services Coordinator, Science and Engineering Libraries

Amy Nurnberger

Research Data Manager, Center for Digital Research and Scholarship (CDRS)

# Outline

1. What is data?
2. Why do we care about data?
3. Why should data be open?
4. How can I make it useful?
5. Organize it
6. Store it
7. Share it
8. Secure it
9. Plan for it

# Outline

1. What is data?
2. Why do we care about data?
3. Why should data be open?
4. How can I make it useful?
5. Organize it
6. Store it
7. Share it
8. Secure it
9. Plan for it

# What is data?

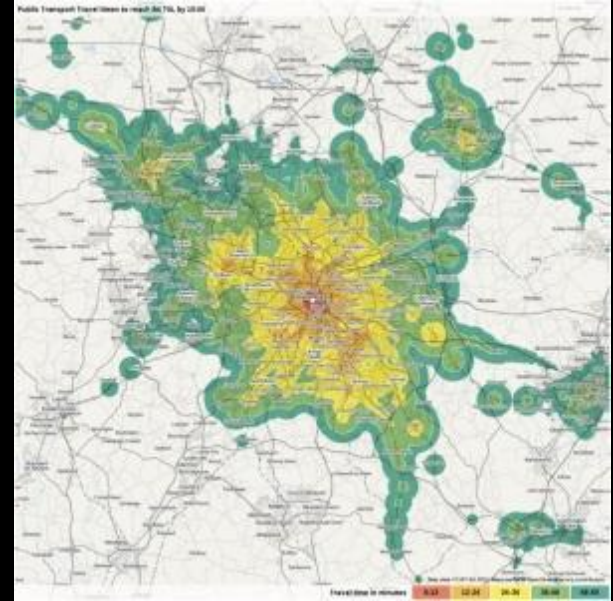


... or is it who?

# What is data?



A basic definition of data is "a collection of observations."



# Outline

1. What is data?
2. Why do we care about data?
3. Why should data be open?
4. How can I make it useful?
5. Organize it
6. Store it
7. Share it
8. Secure it
9. Plan for it

# Why do we care about data?

accurate | credible | reproducible

# Outline

1. What is data?
2. Why do we care about data?
3. Why should data be open?
4. How can I make it useful?
5. Organize it
6. Store it
7. Share it
8. Secure it
9. Plan for it



# Why should data be open?

Data should be open because it ...

- increases the visibility of your research
- may be required
- facilitates new discoveries
- supports open access

# Outline

1. What is data?
2. Why do we care about data?
3. Why should data be open?
4. How can I make it useful?
5. Organize it
6. Store it
7. Share it
8. Secure it
9. Plan for it

# How can I make it useful?

You can:

- organize it
- share it
- secure it
- and plan for it

# Activity 1

1. Pair up
2. First person: open a folder on the second's computer
3. Second person: explain what each file is based on the file name alone!

# Outline

1. What is data?
2. Why do we care about data?
3. Why should data be open?
4. How can I make it useful?
5. Organize it
6. Store it
7. Share it
8. Secure it
9. Plan for it

# Organize it

use meaningful file names

Experiment1.txt

vs.

Cold\_Fusion\_Experiment 2014\_04\_10 EHR.txt

# Organize it

use meaningful file names

Experiment1.txt

vs.



Cold\_Fusion\_Experiment 2014\_04\_10 EHR.txt

# Organize it

use meaningful variable names

var1

vs.

ageGroup



# Organize it

use meaningful variable names

var1

vs.

ageGroup



# Organize it

use meaningful variable names

var1

vs.

ageGroup



Provide a “readme” file or codebook

# Organize it

provide metadata for your data files

- for example:

author/creator

date

file type

data type

software required

# Organize it

use a reference manager to organize your PDFs

- Mendeley and Zotero are great (and FREE!) tools

# Outline

1. What is data?
2. Why do we care about data?
3. Why should data be open?
4. How can I make it useful?
5. Organize it
6. Store it
7. Share it
8. Secure it
9. Plan for it

# Store it

Working space:

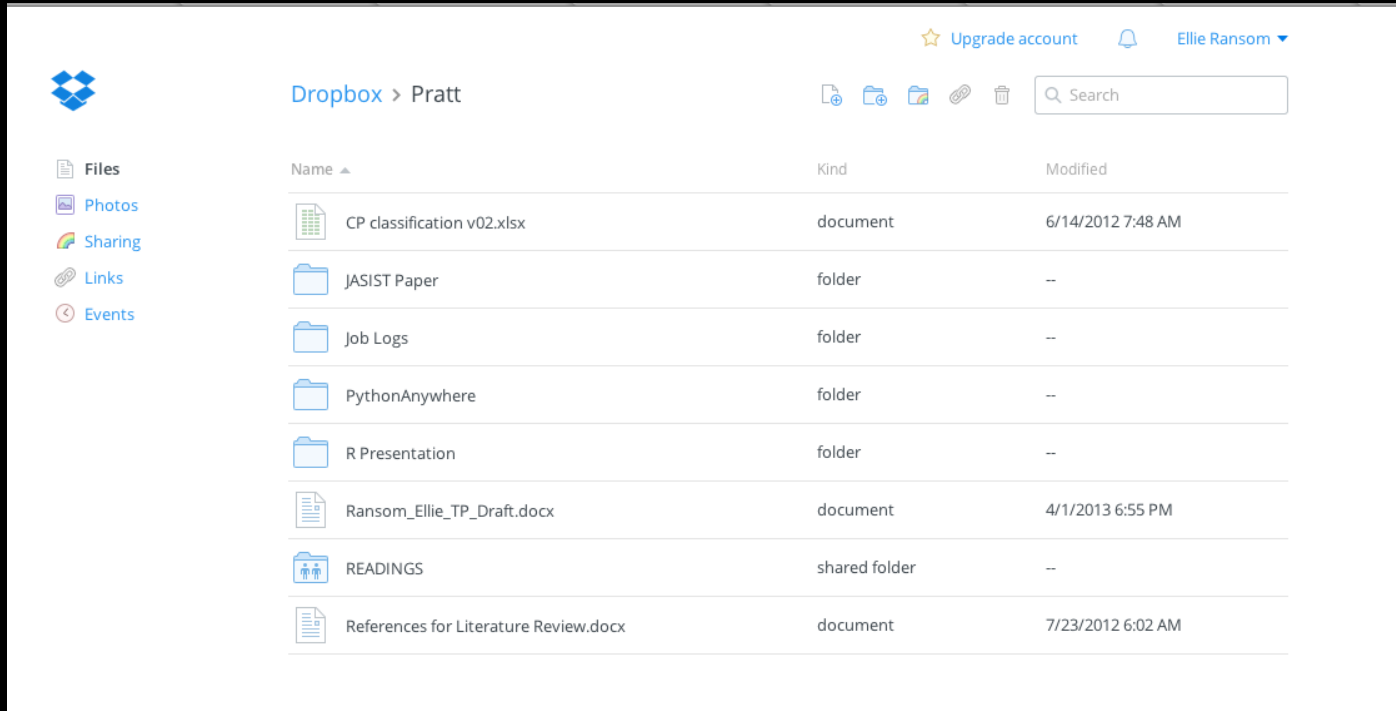
- store in a useful place
- back it up!

we suggest having it in 3 places:

1. hard drive
2. removable hard drive
3. cloud (e.g., github, figshare, dropbox)

# Store it

Dopbox - [www.dropbox.com](http://www.dropbox.com)



The screenshot displays the Dropbox web interface. On the left is a sidebar with navigation links: Files, Photos, Sharing, Links, and Events. The main area shows the 'Pratt' folder. At the top right, there are links for 'Upgrade account' and the user's name 'Ellie Ransom'. Below these are icons for file operations and a search bar. The file list is organized into three columns: Name, Kind, and Modified. The files listed are:

Name	Kind	Modified
CP classification v02.xlsx	document	6/14/2012 7:48 AM
JASIST Paper	folder	--
Job Logs	folder	--
PythonAnywhere	folder	--
R Presentation	folder	--
Ransom_Ellie_TP_Draft.docx	document	4/1/2013 6:55 PM
READINGS	shared folder	--
References for Literature Review.docx	document	7/23/2012 6:02 AM

# Store it

## GitHub - github.com

The screenshot shows the GitHub web interface for a user named 'jeffreylancaster'. At the top, there's a navigation bar with a search bar, links for 'Explore', 'Gist', 'Blog', and 'Help', and the user's profile. Below this is a 'ProTip™' banner. The main content area features a 'News Feed' tab, which is currently selected. Below the tabs is a 'GitHub Bootcamp' section with four numbered steps: 1. Set up Git, 2. Create repositories, 3. Fork repositories, and 4. Be social. Each step includes an illustration and a brief description. Below the bootcamp section, there are two commit notifications from 'SteveRidout' pushed to the 'master' branch of the 'citation-style-editor/csl-...' repository. On the right side, there's a 'Stay in sync with GitHub for Windows' notification and a 'Your repositories (0)' section with a 'New repository' button. The 'Your repositories' section indicates that the user doesn't have any repositories yet and provides links to create a first repository or learn more about Git and GitHub.

jeffreylancaster Search or type a command Explore Gist Blog Help jeffreylancaster

ProTip™ You can get a list of all commits a branch has pointed to with `git reflog` News Feed

News Feed Pull Requests Issues Stars

**GitHub Bootcamp** If you are still new to things, we've provided a few walkthroughs to get you started.

- Set up Git**  
A quick guide to help you get started with Git.
- Create repositories**  
Repositories are where you'll work and collaborate on projects.
- Fork repositories**  
Forking creates a new, unique project from an existing one.
- Be social**  
Send pull requests, follow friends. Star and watch projects.

September 13, 2013  
SteveRidout pushed to master at citation-style-editor/csl-...  
7820ec2 require build directory argument

September 13, 2013  
SteveRidout pushed to master at citation-style-editor/csl-...  
a8a9104 update readme

September 13, 2013

Stay in sync with GitHub for Windows  
Quickly see when a branch is no longer in sync in GitHub for Windows

x hide this broadcast View 85 new broadcasts

Your repositories (0) New repository

You don't have any repositories yet!  
Create your first repository or learn more about Git and GitHub.



# Store it

FigShare - figshare.com



## 1GB of private space

taggable and easily filtered, your research data is better managed and easy to locate



## Upload all formats

all research outputs welcomed – images, graphs, videos, datasets



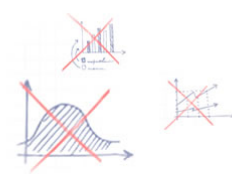
## Unlimited public space

upload to your heart's content  
the more - the better



## Quick & simple upload

even your PI can use it!



## Publish negative data

all published research is citable



## Cloud based

Secure and accessible from anywhere

# Store it

## SlideShare - www.slideshare.net

The screenshot displays the SlideShare website interface. At the top, there is a navigation bar with the SlideShare logo, an 'Explore' dropdown menu, a search bar, and links for 'Upload', 'Login', and 'Signup'. Below the navigation bar, a banner area features the text 'Discover. Share. Present.' and 'Share what you know and love through presentations, infographics, documents and more.' A section titled 'Today's Top SlideShares' showcases three featured presentations: 'Doing Business in China FUNDAMENTALS' by Pengyuan Zhao (1,652 views), 'Turn On Experience' by Jack Morton Worldwide (5,041 views), and 'The literary giant Gabriel Garcia Marquez' by guimera (1,416 views). Below this, a 'Featured SlideShares' section displays a grid of various presentations, including 'Starting from the Ground UP' by Judy Wang, '5 Smart Ways to Thank Administrative Professional' by Cotap, '27 Strategies for Breaking ANY Bad Habit' by Steve Scott, 'Implementation of systems for Media / Digital Asset Management in' by Kara Van Malssen, '5 Brilliant short books' by EHSB+, and 'How I Learned to Stop Worrying and Love Google GLASS' by Wolf Loescher. The interface also includes a 'See all >>' link and an 'Ads by LinkedIn' section.

# Store it

Archived space:

- subject specific repository
- Academic Commons - CU's institutional repository

# Outline

1. What is data?
2. Why do we care about data?
3. Why should data be open?
4. How can I make it useful?
5. Organize it
6. Store it
7. [Share it](#)
8. Secure it
9. Plan for it

# Share it

# Share it

## Scholarly Communication



Research Without Borders:  
Open Access in the  
Americas



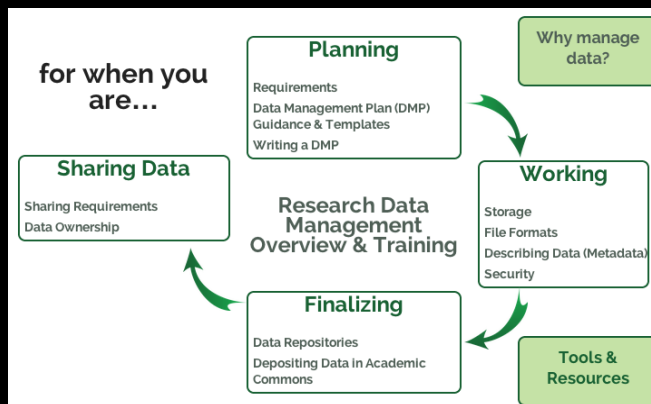
# Share it

## Scholarly Communication

## Research Data Management

Research Without Borders:  
Open Access in the  
Americas

OPEN  ACCESS



# Share it

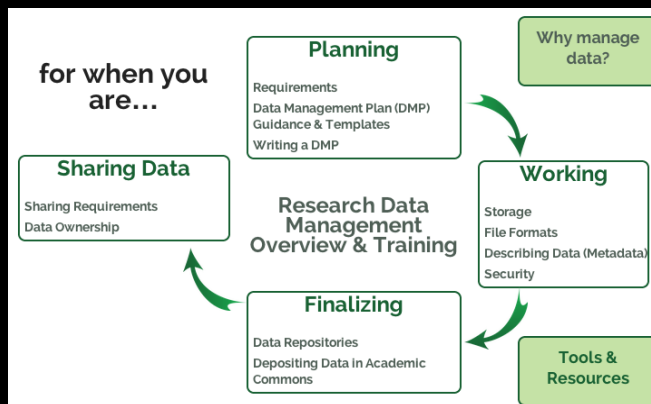
## Scholarly Communication

## Research Data Management

## Publications

Research Without Borders:  
Open Access in the  
Americas

OPEN  ACCESS



**Tremor**  
and Other Hyperkinetic Movements

**cusj**  
COLUMBIA  
UNDERGRADUATE  
SCIENCE JOURNAL



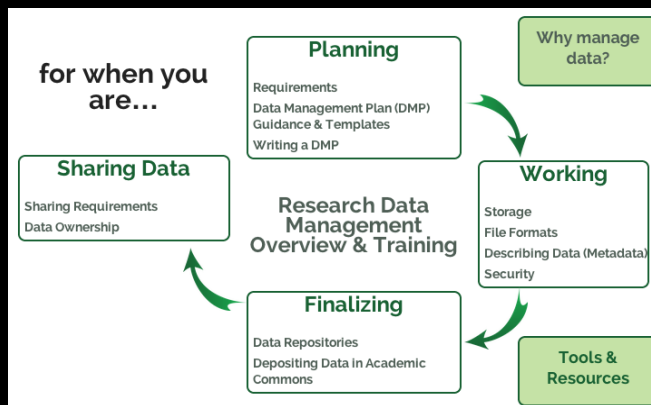
# Share it

## Scholarly Communication

Research Without Borders:  
Open Access in the  
Americas

OPEN  ACCESS

## Research Data Management



## Publications

**Tremor**  
and Other Hyperkinetic Movements

**cusj**  
COLUMBIA UNDERGRADUATE  
SCIENCE JOURNAL

## Institutional Research Repository

 COLUMBIA UNIVERSITY  
**ACADEMIC COMMONS**

# Share it

## Scholarly Communication

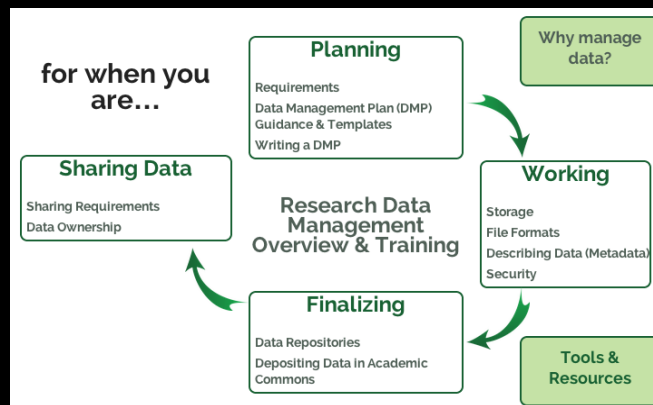
## Research Data Management

## Publications

## Institutional Research Repository

Research Without Borders:  
Open Access in the  
Americas

OPEN  ACCESS



<http://cdrs.columbia.edu/>  
<http://scholcomm.columbia.edu/>  
<http://scholcomm.columbia.edu/data-management/>  
<http://cdrs.columbia.edu/cdrsmain/publicationsservices/>  
<http://academiccommons.columbia.edu/>  
[CU-RDM@columbia.edu](mailto:CU-RDM@columbia.edu)

**Tremor**  
and Other Hyperkinetic Movements

**cusj**  
COLUMBIA UNDERGRADUATE SCIENCE JOURNAL

 COLUMBIA UNIVERSITY  
**ACADEMIC COMMONS**

# Share it



The screenshot shows the Columbia University Academic Commons homepage. At the top is a dark blue header with the university's crest and the text "COLUMBIA UNIVERSITY ACADEMIC COMMONS". Below the header is a search bar with a magnifying glass icon and a "Search" button. A large image of a building at night is on the left. To its right is a paragraph describing the repository. The main content area is divided into several sections: "Log Out [aln2141] | Saved Searches", "BROWSE AC CONTENT" with sub-sections for "DEPARTMENTS" and "SUBJECTS", "HELPFUL LINKS" with links to "DEPOSIT YOUR RESEARCH", "ABOUT ACADEMIC COMMONS", and "FAQ/ASK A QUESTION", "AC BY THE NUMBERS" showing statistics like "11216 items in Academic Commons" and "3080 Objects added in the last year", and "DEPOSIT YOUR WORK IN ACADEMIC COMMONS". A central section titled "New in Academic Commons:" features three highlighted items: "Two Papers of Financial Engineering Relating to the Risk of the 2007--2008 Financial Crisis" by Zhong, Haowen; "RDA Plenary Twitter Archive, September, 2013" by Nurnberger, Amy L.; and "Exploring a behavioral role for presynaptic inhibition at spinal sensory-motor synapses" by Fink, Andrew. At the bottom are two more items: "Kernel-based association measures" and "Notch Signaling Determines Lymphatic Cell Fate and".

## Preserved

Backed up & backed up again  
Checksums

## Findable

Search engine optimization  
Digital object identifiers (doi)

## Shareable



CENTER FOR DIGITAL RESEARCH  
AND SCHOLARSHIP

COLUMBIA UNIVERSITY LIBRARIES/INFORMATION SERVICES

# Activity 2

A list of subject repositories is provided, find your subject and look up what it takes to deposit a data set into the repository.

# Activity 2

A list of subject repositories is provided, find your subject and look up what it takes to deposit a data set into the repository.

Medical, Biomedical – PubMed (NIH)

<http://www.ncbi.nlm.nih.gov/guide/howto/submit-data/>

Mathematics and Statistics; Physics and Astronomy;

Computers and IT – arXiv

<http://arxiv.org/help/submit>

Life Sciences – bioRxiv

<http://biorxiv.org/submit-a-manuscript>

Computer Science – GitHub

<https://github.com/>

Earth & Environmental – EarthChem

<http://www.earthchem.org/data/contribute>

Can't find yours?

General Subject Matter - Academic Commons

<http://academiccommons.columbia.edu/>

Know of one we don't? Use that one!

# Share it

... on social media



“Research not communicated is research not done”

Prof. Anne Glover, EU Science Advisor

# Share it

Sign up for an  
ORCID

# Share it

Sign up for an  
ORCID

John Smith, John D. Smith, J.  
Smith, J.D. Smith



# Share it

Sign up for an  
ORCID

John Smith, John D. Smith, J.  
Smith, J.D. Smith

L. Wang, Li Wang, Wang Li, W. Li

# Outline

1. What is data?
2. Why do we care about data?
3. Why should data be open?
4. How can I make it useful?
5. Organize it
6. Store it
7. Share it
8. Secure it
9. Plan for it

# Secure it

dealing with sensitive data? make sure it is secure.

does your research fall under HIPAA or FERPA guidelines?

# Outline

1. What is data?
2. Why do we care about data?
3. Why should data be open?
4. How can I make it useful?
5. Organize it
6. Store it
7. Share it
8. Secure it
9. Plan for it

# Plan for it

Data Management Plans - does your grant require one?

[scholcomm.columbia.edu](https://scholcomm.columbia.edu)

# Homework

**Researchers: Ask your PI for a copy of an accepted grant**

**PI: Show your research staff the accepted grant they are working on**

# **Test out Electronic Lab Notebooks**

[http://library.columbia.edu/locations/dsc/elns\\_trial.html](http://library.columbia.edu/locations/dsc/elns_trial.html)



# Instagram

1. Take a picture with this month's theme.
2. Tag @columbiascience.
3. Win prizes!

April theme: **#experiment**

more details at: [bit.ly/instagramScience](http://bit.ly/instagramScience)



# Questions?

Ellie Ransom

Research Services Coordinator,  
Science and Engineering Libraries

[ehr2125@columbia.edu](mailto:ehr2125@columbia.edu)

Amy Nurnberger

Research Data Manager,  
Center for Digital Research and  
Scholarship (CDRS)

[CU-RDM@columbia.edu](mailto:CU-RDM@columbia.edu)