

A wooden smiley face with a red heart on its cheek and a sign that says "WANT TO GO ON A DATA?". The sign is orange with black text. The smiley face is made of wood and has two small wooden dots for eyes. The background is dark.

# NSF and Beyond: Research data management for grant administrators

Dorothea Salo  
Jan Cheetham  
NCURA Region IV Meeting  
Spring 2013

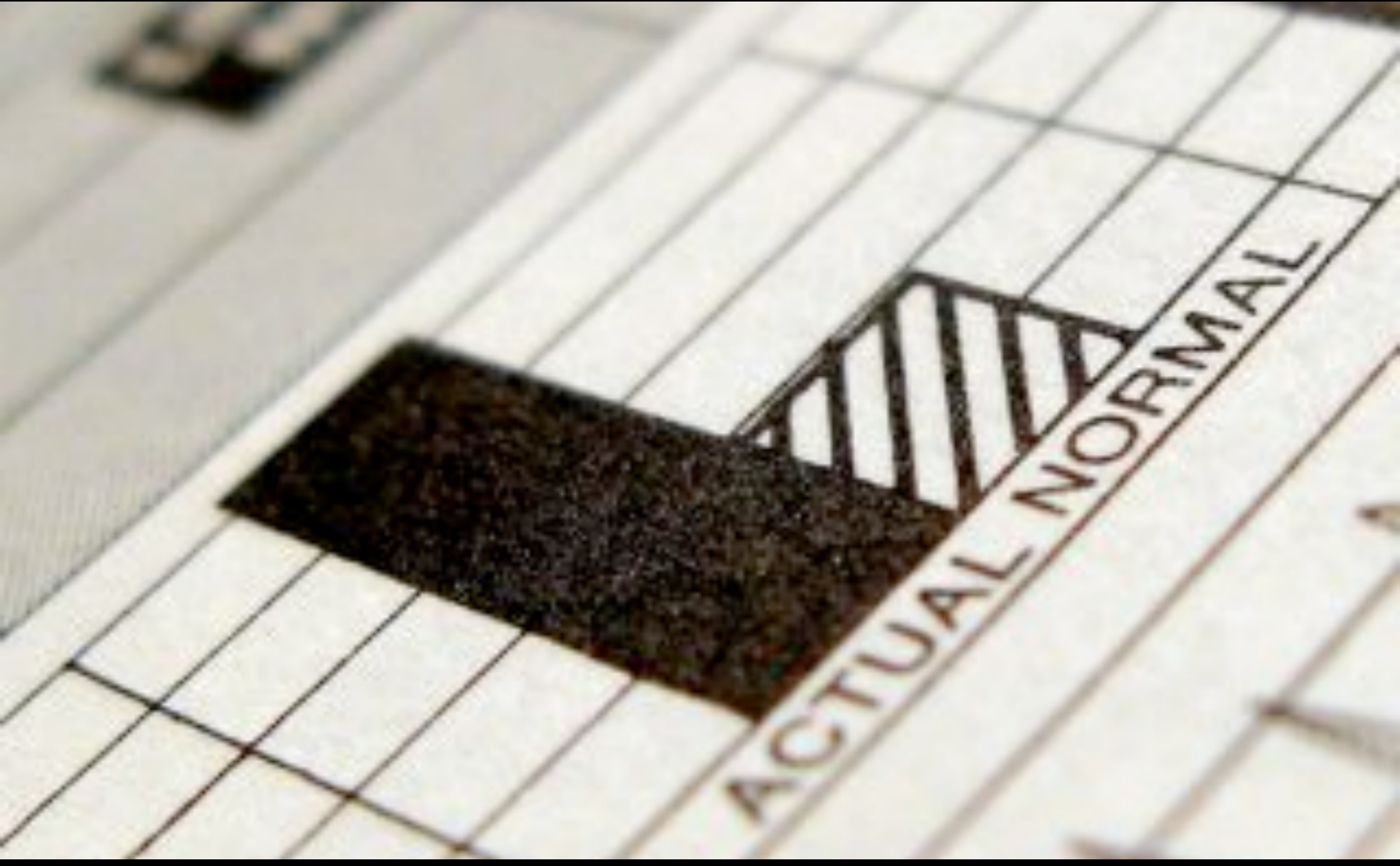
When many people think



Data!

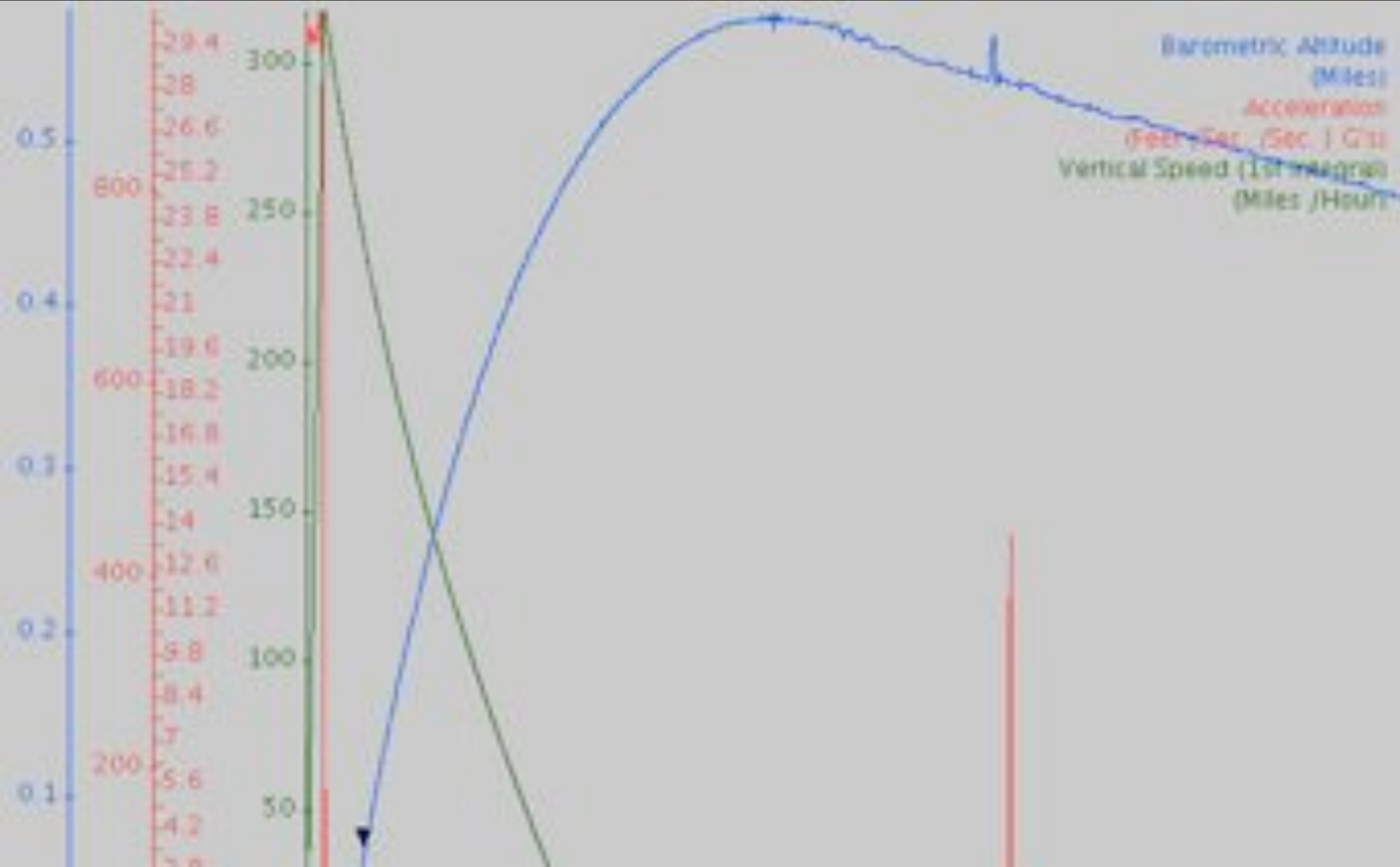
# chances are they think

Photo: kevin dooley, "Actual is not normal (a tribute to Edward Tufte)" <http://www.flickr.com/photos/pagedooley/2121472112/> CC-BY



# or

Photo: Steve Jurvetson, "Rocket Flight Computer Readout" <http://www.flickr.com/photos/jurvetson/4093645514/> CC-BY



# or even

Photo: T Farrant of GDS Infographics, "Ballooning CEO Salaries and Mass Layoffs" <http://www.flickr.com/photos/gdsdigital/4963409391/> CC-BY

## BALLOONING CEO SALARIES AND MASS LAYOFFS



# Data trapped in amber

Photo: Luz, "amber" <http://www.flickr.com/photos/nieve44/3800137286/> CC-BY



# Data trapped in notebooks

Photo: Steve Jurvetson, "How the Eagle Landed – the Grumman Construction Log" <http://www.flickr.com/photos/jurvetson/7610058658/> CC-BY



But that's **changing!**

(you know, just like  
everything else, right?)



# Research is changing...

- Research datasets were second-class citizens.
  - Publications were all that mattered!
  - And publishing data in print was uneconomical even when possible.
  - So nobody saw anybody's data.
- Data are now digital. The game changes!
  - Data are shared more, and more openly! Open Source, Open Access, Open Data.
  - There's a lot still to be worked out about how to share, cite, credit, and license digital data.
- This is not going away. The emphasis on data as a first-class research product will only intensify.

# What's in it for researchers?

- Additional trust in their publications.
- Additional citations FOR their publications.
- Additional collaborators
  - including across disciplines
- Additional publications, possibly...
  - ... if they publish in a data or code journal.

# Funder policies



The image is a screenshot of the National Science Foundation (NSF) website. At the top left is the NSF logo, a blue globe with 'NSF' in white, and the text 'National Science Foundation' and 'WHERE DISCOVERIES BEGIN' below it. To the right is a search bar with the word 'SEARCH' and a magnifying glass icon. Above the search bar is an orange button labeled 'QUICK LINKS'. Below the search bar is a dark blue navigation bar with white text for 'FUNDING', 'AWARDS', 'DISCOVERIES', 'NEWS', 'PUBLICATIONS', 'STATISTICS', 'ABOUT NSF', and 'FASTLANE'. On the left side, there is a sidebar with a black box containing the text 'Office of Budget, Finance and Award Management (BFA)' and a green and yellow abstract image. Below this are links for 'DIAS Home', 'CAAR Branch', 'Policy Office', 'Systems Office', and 'View DIAS Staff'. At the bottom of the sidebar is a search box for 'Search DIAS Staff'. The main content area has a large blue heading 'Dissemination and Sharing of Research Results'. Below this is a sub-heading 'NSF Data Sharing Policy' followed by a paragraph: 'Investigators are expected to share with other researchers, at no more than incremental cost and within a reasonable time, the primary data, samples, physical collections and other supporting materials created or gathered in the course of work under NSF grants. Grantees are expected to encourage and facilitate such sharing. See [Award & Administration Guide \(AAG\) Chapter V.D.4.](#)' Below this is another sub-heading 'NSF Data Management Plan Requirements' followed by a paragraph: 'Proposals submitted or due on or after January 18, 2011, must include a supplementary document of no more than two pages labeled "Data Management Plan". This supplementary document should describe how the proposal will conform to NSF policy on the dissemination and sharing of research results. See [Grant Proposal Guide \(GPG\) Chapter II.C.2.1](#) for full policy implementation.'

# OSTP Public Access Memo

Access to digital data sets resulting from federally funded research allows companies to focus resources and efforts on understanding and exploiting discoveries. For example, open weather data underpins the forecasting industry, and making genome sequences publicly available has spawned many biotechnology innovations. In addition, wider availability of peer-reviewed publications and scientific data in digital formats will create innovative economic markets for services related to curation, preservation, analysis, and visualization. Policies that mobilize these publications and data for re-use through preservation and broader public access also maximize the impact and accountability of the Federal research investment. These policies will accelerate scientific breakthroughs and innovation, promote entrepreneurship, and enhance economic growth and job creation.

# OSTP Public Access Memo

Further, each agency plan for both scientific publications and digital scientific data must contain the following elements:

- a) a strategy for leveraging existing archives, where appropriate, and fostering public-private partnerships with scientific journals relevant to the agency's research;
- b) a strategy for improving the public's ability to locate and access digital data resulting from federally funded scientific research;

# Who else?

- Alphabet soup! NIH, NOAA, NEH ODP, DOE, EPA, IMLS...
- Great source for federal data policies: <http://www.cdlib.org/services/uc3/datamanagement/funding.html>
- Private funders, e.g. Moore Foundation, World Bank
- SHERPA/JULIET: <http://www.sherpa.ac.uk/juliet/>
  - Early days, but this is still a useful resource!

# Institutional data policies



The Graduate School

Policies and Procedures

Conflict of Interest

Conflict of Interest  
Regulatory Changes

Institutional COI

Research Data & Tangible  
Research Property Policies

Ethical Principles, Federal  
& State Law

Export Control

Extramural Support  
Policies

Human Research  
Protection Program

PI Status

Outside Activities  
Reporting

## UW-Madison Policies on Research Data & Tangible Research Property Policies

### Policy on Data Stewardship, Access and Retention

The University of Wisconsin-Madison has established this policy on Data Stewardship, Access and Retention to assure that research data are appropriately maintained, archived for a reasonable period of time, and available for review and use under the appropriate circumstances. The policy also provides for transfer of data in the event a research leaves UW-Madison.

This policy applies to all University of Wisconsin-Madison faculty, academic staff, visiting scholars, postdoctoral fellows or other trainees, research technicians, and graduate or undergraduate students and any other persons at UW-Madison involved in the design, conduct or reporting of research at or under the auspices of UW-Madison, and it applies to all research projects on which those individuals work, regardless of the source of funding for the project.

# Journal data policies

JISC

Search for  in

[Home](#) | [About Jisc](#) | [Supporting your institution](#) | [Projects, programmes & services](#) | [Funding](#) | [Publications](#) | [Blog](#)

## ▶ Activities by Topic

Programmes

Digital infrastructure:  
Research management  
programme

Managing research  
data

Innovative Research  
Data Publication

• Journal Research  
Data Policy Bank  
(JoRD)

[Home](#) • [Projects, programmes & services](#) • ... • [Innovative Research Data Publication](#) • [Journal Research Data Policy Bank \(JoRD\)](#)

## Journal Research Data Policy Bank (JoRD)

The Journal Research Data Policy Bank (JoRD) project will be conducting a feasibility study into the scope and shape of a sustainable service to collate and summarise journal policies on Research Data. The aim of this service will be to provide researchers, managers of research data and other stakeholders with an easy source of reference to understand and comply with Research Data policies.

Through maintaining a firm focus upon research literature and stakeholder consultations, this project will identify and consult with a wide range of stakeholders (nationally and internationally) and look at journal policies on Research Data from leading journals to deliver detailed requirements and



# Elements of a DMP

	General	Specific
Example	NSF-BIO NSF-ENG NEH	NOAA Some NSF solicitations
Description of data/ metadata	Variable, community standards evolving	Specific formats may be required
Preservation/Retention	Cyber resources for storage and backup	Plans for data lifecycle, including long-term curation
Access and Sharing	Timing, methods for sharing, policies for re-use	Timing, repositories, communities may be specified

# What help do researchers need?

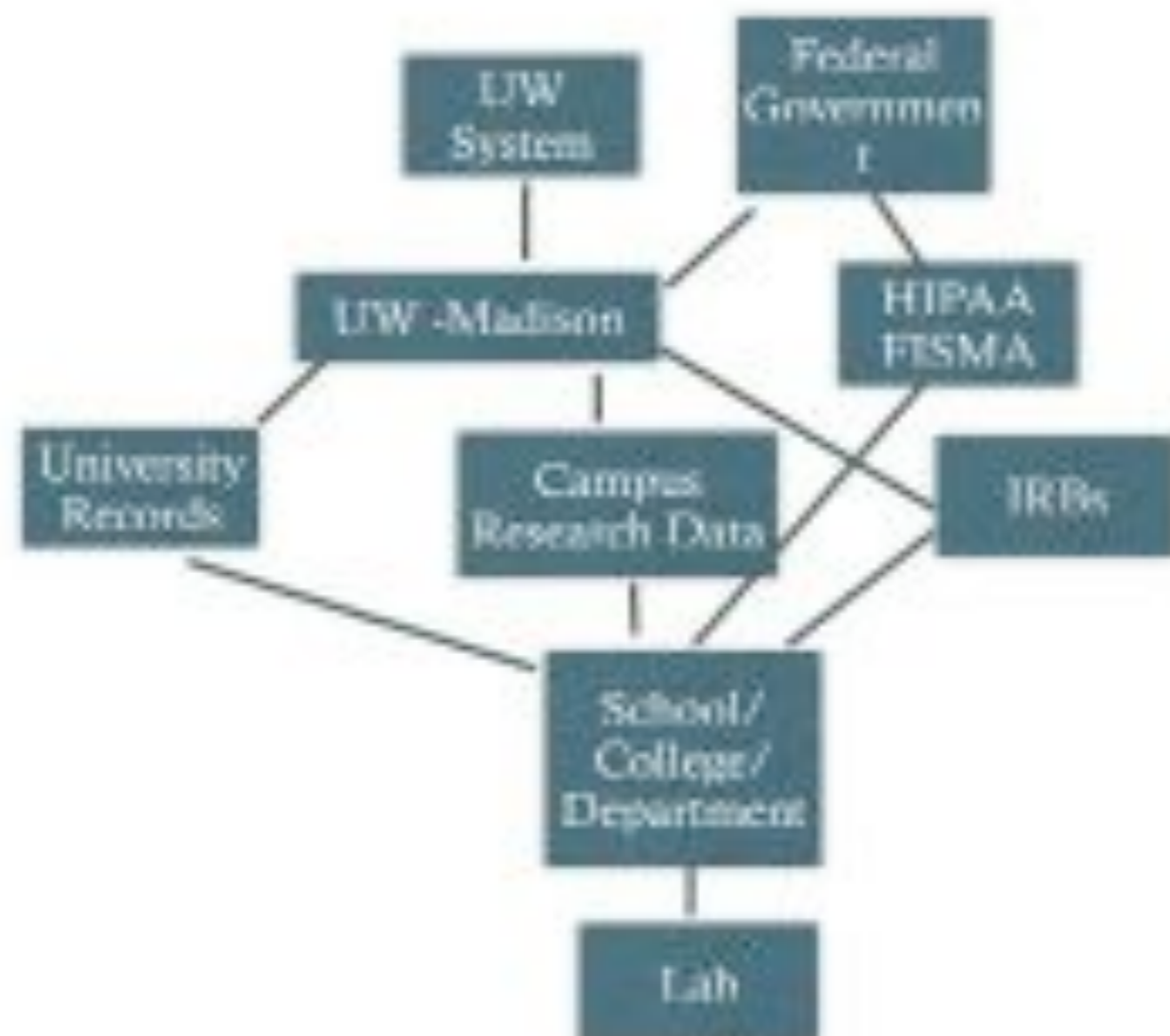


# Basic storage and preservation

- Storage and backup on maintained campus servers
- Long-term preservation



# Data Policies



# Documentation

I.e., metadata, can be highly variable

- Chemistry—what's recorded in the lab notebook
- Many life and social sciences areas---“read me” document describing text or spreadsheet data
- Earth, Environmental, Physical Sciences—generally well-accepted standards and formats

A screenshot of a spreadsheet application. The spreadsheet has a grid of cells with various colors (green, blue, red, yellow, grey) applied to different columns. The data is organized into a table with multiple rows and columns. The interface includes a menu bar at the top and a status bar at the bottom.

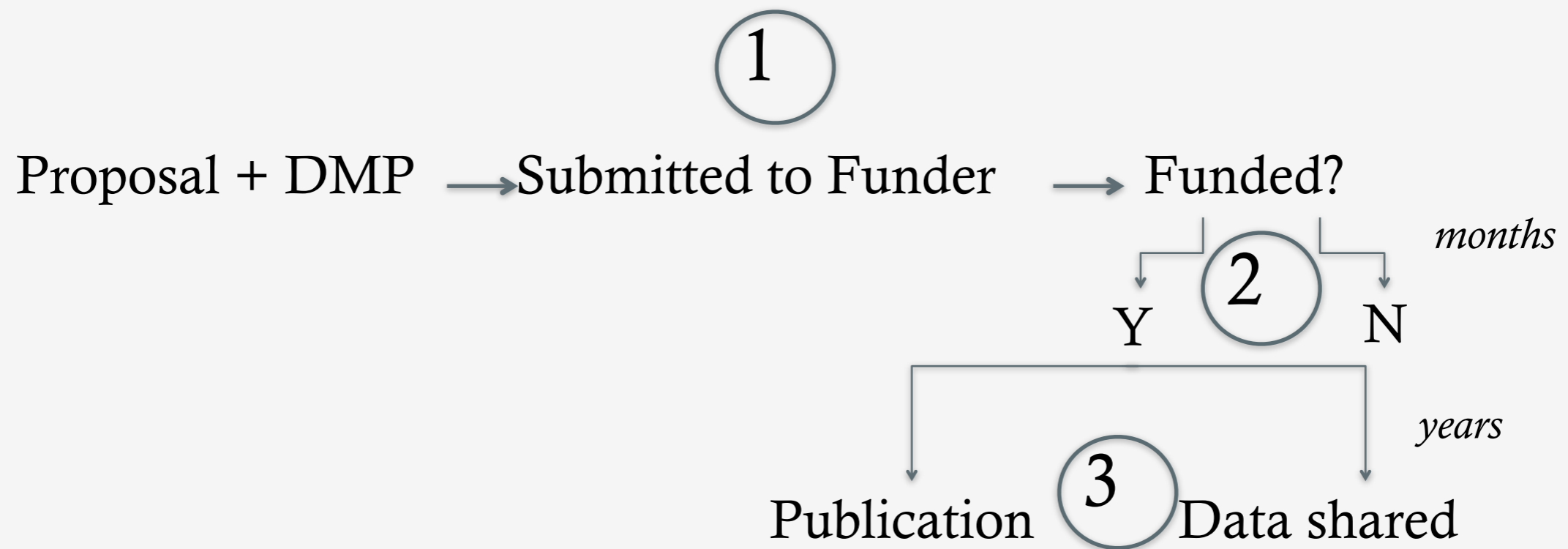
# Data Sharing

- Data sharing
  - Publications
  - Websites
  - Public repositories
- Documentation to make data understandable
  - Standards in some disciplines
  - Others emerging

## GEOGRAPHIC OVERLAP AREAS (USGS)

```
shp2pgsql -s 4269  
countyp020 county  
>county.sql  
shp2pgsql -s 4269  
hucs00p020 huc  
>huc.sql  
shp2pgsql -s 4269  
quad24 quad  
>quad.sql
```

# Is help with DMPs working?





What are we doing?



# Who's "we," anyway?

- Campus research offices, research-policy groups
- Campus IT organizations
  - both formal and grassroots
- Libraries, librarians, library schools
- Disciplinary societies and data centers
- Web startups and other third-party providers

# Here's us (Madison)

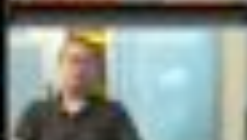


## Research Data Services

DIGITAL CURATION AT THE UNIVERSITY OF WISCONSIN-MADISON

Writing a Data Plan ▾ Managing Data ▾ Sharing Data ▾ Events ▾ About ▾

DATAVISUALIZATION



# DATAVISUALIZATION

Seeing Everything and Making Sense



April 25 1:30-2:45 • Ebling Symposium Ctr, Rm1220 - Microbial Sciences Bldg

1000 5000  
ORGANIZING COLLECTING  
STORING PRESERVING  
SHARING  
ANALYZING  
VISUALIZING  
ARCHIVING



We **love** your data too.

### News

- Tentative schedule for OR meeting announced
- [CZO-NO] Program for Submission in – National Science Foundation
- Artificial Retina Receives FDA Approval – National Science Foundation

[More news >>](#)

### Events

Today ◀ ▶

Thursday, April 11

Shared Horizons: Data, Biomedicine

Friday, April 12

Shared Horizons: Data, Biomedicine

Tuesday, April 16

12:00pm Technology and Data of th

Tuesday, May 7

12:00pm Tracking Behavioral Data:

Sunday, May 12

DigCCur Professional Institute: Cur

Monday, May 13

DigCCur Professional Institute: Cur

Tuesday, May 14

# Here's Milwaukee!



## Graduate School

Advancing Research, Scholarship, & Innovation



[Prospective & Current Graduate Students](#)

[UWM Research](#)

[Faculty & Staff Resources](#)

[Corporate Collaborations & Technology Transfer](#)

[Contact Us](#)

[UWM Home](#) • [Graduate School](#) • [Research](#) • [Researcher Central](#) • [Developing a Proposal](#) • [Data Management](#)

*Developing a Proposal*

## Data Management Plan

POWERFUL IDEAS | PROVEN RESULTS

### Researcher Central

[Developing Research Ideas](#)

[Finding Funding](#)

[Developing a Proposal](#)

[Submitting a Proposal](#)

[Forms](#)

[Developing International Research Projects](#)

[Managing Grants](#)



### Data Management Plan Defined

A description of the production, organization, preservation, or accessing of data a project will produce or gather.



### Data Defined

There are four categories defined by the federal Office of Management and Budget (OMB).



### Ontologies

Conceptualizations of domains of information that show the categories of "things" in the domain.



### Metadata

information about the context, content, quality, provenance, and/or accessibility of a set of data.



### Data File Formats

# Capturing and managing data



An open source tool helping researchers document, manage, and archive their tabular data, DataUp operates within the scientist's workflow and integrates with Microsoft® Excel.

DataUp Features

Start Using DataUp

Customize DataUp

Contact Us

What  
**DataUp**  
can do for you

- **Check for Best Practices**
- Create Metadata
- Get Credit
- Archive & Share

## Check for Best Practices

Find out if characteristics of your dataset will prevent its future use.

SEE MORE

# Sharing and storing data



Browse

Upload



get credit for all your research

or store it privately for FREE\*

Feedback?

# Keeping data safe

**PURDUE**  
UNIVERSITY

**Purdue University Research Repository**

Login

Register

Report a bug

[Home](#) [Resources](#) [Projects](#) [Get Started](#) [Contact Us](#)

Search

## Publish Datasets with DOIs

Use PURR to publish datasets with Digital Object Identifiers (DOI) that make it easier for people to cite your data and give you credit. Purdue is a founding member of DataCite, the international agency that registers DOIs for data.

[Learn More](#)



DataCite

<http://dx.doi.org/10.4231/D3992W550>



## Start Your Research Project

Create a Data Management Plan

## Featured Dataset

Graph of Flickr Photo-Sharing Social Network

## Do you have a question?

Ask a Librarian

# Helping people find data



## Databib

[Find Repositories](#) | [Submit](#) | [Connect](#) | [About](#)

[Login/Regi...](#)

**Databib** is a searchable catalog / registry / directory / bibliography of **research data repositories**.

Search

[Find](#)

[Advanced Search](#)

Browse [ [Subjects](#) | [A](#) [B](#) [C](#) [D](#) [E](#) [F](#) [G](#) [H](#) [I](#) [J](#) [K](#) [L](#) [M](#) [N](#) [O](#) [P](#) [Q](#) [R](#) [S](#) [T](#) [U](#) [V](#) [W](#) [X](#) [Y](#) [Z](#) | [All](#) ]

**3**

**3TU.Datacentrum**

A multidisciplinary data repository for a consortium of universities in the Netherlands housing over...

**A**

**Addgene Plasmid Database**

Addgene is a non-profit organization dedicated to making it easier for scientists to share plasmids....

**Adult Blood Lead Epidemiology and Surveillance (ABLES) Interactive Database**

ABLES provides data on lead exposure of adults in the United States. The data comes from laboratory...

**Advanced Cooperative Arctic Data and Information Service (ACADIS)**

The Advanced Cooperative Arctic Data and Information Service (ACADIS) program includes data manageme...

### Recently added...

[Earth System Research Laboratory \(ESRL\): Global Monitoring Division](#)

[BioGrid Australia Limited](#)

[LAADS Web \(Level 1 and Atmosphere Archive and Distribution System\)](#)

[Integrated Fertility Survey Series \(IFSS\)](#)

[Advanced Cooperative Arctic Data and Information Service \(ACADIS\)](#)

[Twitter](#)



515 data repositories total in Databib.

# Citing data

[Home](#)[Members](#)[FAQs](#)[Services](#)[Resources](#)[Events](#)[Contact us](#)

DataCite

Helping you to find,  
access, and reuse data

## What is DataCite?

We are a not-for-profit organisation formed in London on 1 December 2009.  
Our aim is to:

- establish easier access to research data on the Internet
- increase acceptance of research data as legitimate, citable contributions to the scholarly record
- support data archiving that will permit results to be verified and re-purposed for future study.

These goals are laid down in the DataCite statutes.

Why cite  
data?

What is  
DataCite?

What do  
we do?



# Human infrastructure: education and training

The logo for MANTRA, featuring the word "MANTRA" in a stylized, white, handwritten-style font.

Research Data Management Training

[Home](#) | [Software practicals](#) | [Project overview](#) | [University of Edinburgh guidance](#) | [Testimonials](#) | [Acknowledgements](#) | [Feedback](#)

## Online learning units

-  [Introduction to the course](#)
-  [Research data explained](#)
-  [Data management plans](#)
-  [Organising data](#)
-  [File formats & transformation](#)
-  [Documentation & metadata](#)
-  [Storage & security](#)
-  [Data protection, rights & access](#)  
NEW
-  [Preservation, sharing & licensing](#)

Introduction to the course

## About the course

This is a non-credit, free course which provides guidelines for good practice in research data management.

The course is particularly appropriate for postgraduate students and early career researchers who work with data and would like to learn more about managing their research data.

The course content is mainly geared for three disciplines: geosciences, social and political sciences and clinical psychology, however, many of the issues covered apply equally to all research disciplines.



We need you too.



**What we need from you:  
connections!**

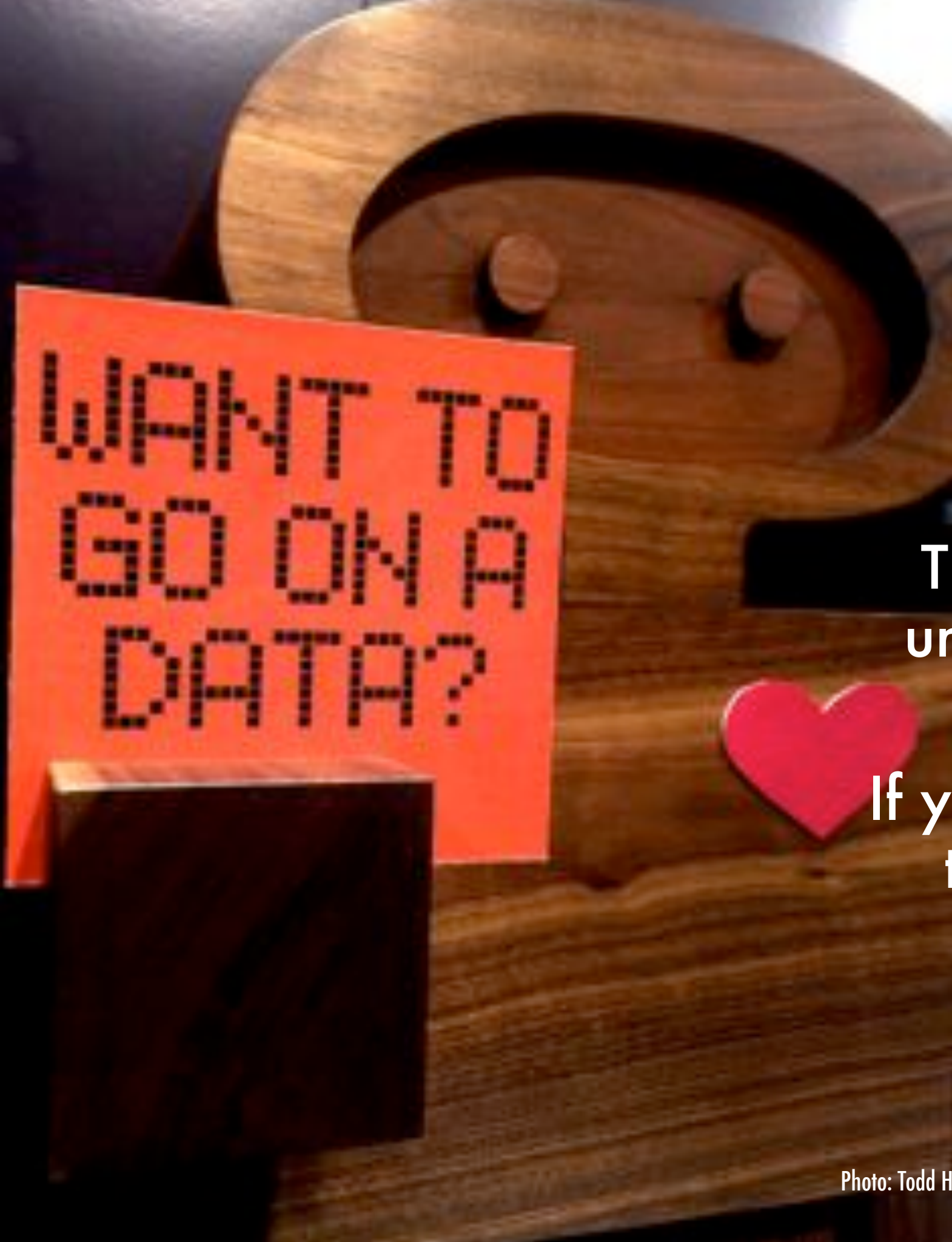
# Benefits, challenges?

Have you seen any immediate wins at your campus?

- Cultural shift about data responsibilities?
- Attendance at data management workshops?

What are the challenges?

- Funding for data management?
  - Storage and backup
  - Tools to automate workflow and capture metadata
- Cultural changes about public access to data?
- Other?

A wooden sculpture of a smiley face with a red heart on its chest and a red sign on its left side. The sign has the text "WANT TO GO ON A DATA?" in a pixelated font.

WANT TO  
GO ON A  
DATA?

# Thank you!

This presentation available  
under a Creative Commons  
3.0 Attribution license.

If you reuse, please continue  
to credit included images!