UCLA
Presentations

Title
Local or global? Making sense of the data sharing imperative (Keynote)

Permalink
https://escholarship.org/uc/item/0m51t83m

Author
Borgman, Christine L.

Publication Date
2011

Copyright Information
This work is made available under the terms of a Creative Commons Attribution-NonCommercial-NoDerivatives License, availalbe at https://creativecommons.org/licenses/by-nc-nd/4.0/

eScholarship.org Powered by the California Digital Library
University of California

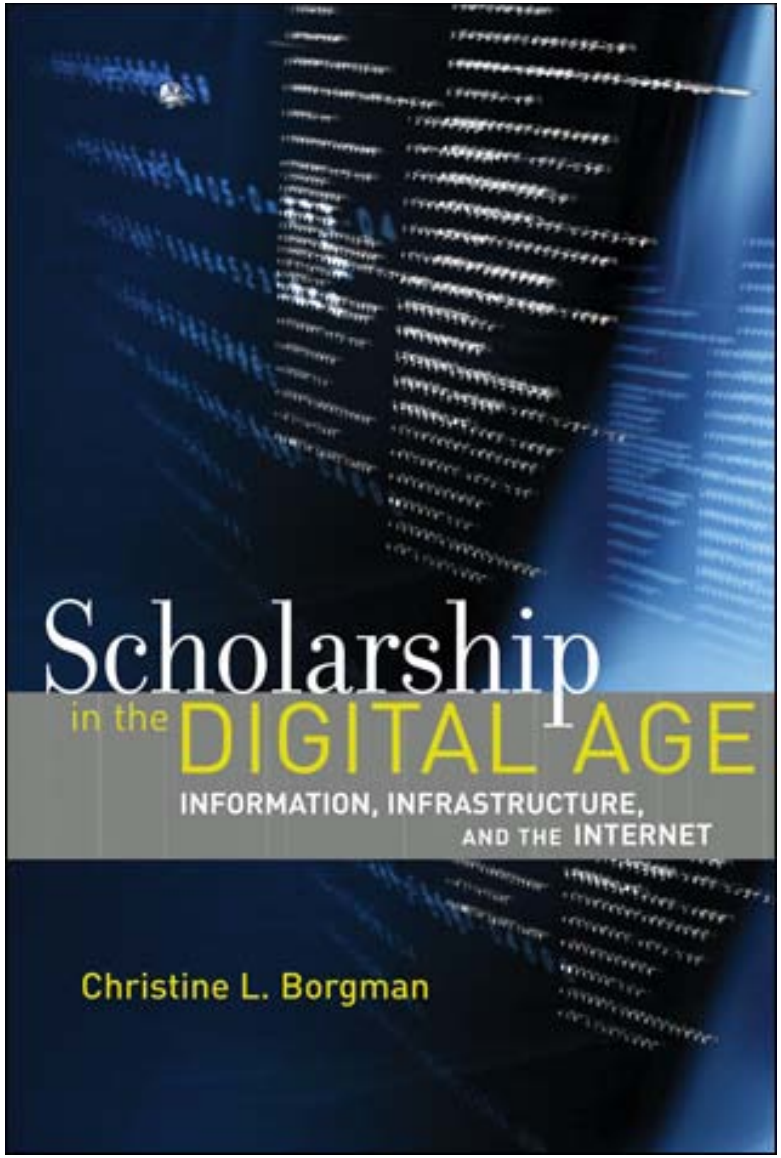
Local or global? Making sense of the data sharing imperative

Christine L. Borgman
Professor & Presidential Chair in Information Studies
University of California, Los Angeles

Keynote Presentation

Oxford Internet Institute

A Decade in Internet Time: Symposium on the Dynamics of the Internet and Society



Deluge!!!

http://www.guzer.com/pictures/suprise_suprise.jpg

Revolutionizing Science and Engineering Through Cyberinfrastructure:

Report of the National Science Foundation Blue-Ribbon Advisory Panel on Cyberinfrastructure

January 2003

Daniel E. Atkins, Chair
University of Michigan

Kelvin K. Droegemeier
University of Oklahoma

Stuart I. Feldman
IBM

Hector Garcia-Molina
Stanford University

Michael L. Klein
University of Pennsylvania

David G. Messerschmitt
University of California at Berkeley

Paul Messina
California Institute of Technology

Jeremiah P. Ostriker
Princeton University

Margaret H. Wright
New York University

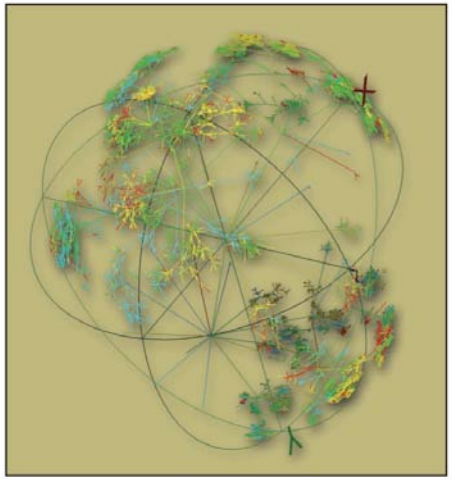
National Science Board

Long-Lived Digital Data Collections: Enabling Research and Education in the 21st Century

National Science Foundation — September 2005

2020 SCIENCE

Cyberinfrastructure Vision for 21st Century Discovery



National Science Foundation
Cyberinfrastructure Council
March 2007

UKOLN

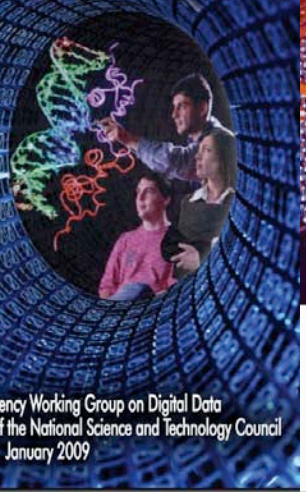
Dealing with Data: Roles, Rights, Responsibilities and Relationships

Consultancy Report

Document details

Author:	Dr Liz Lyon, UKOLN, University of Bath
Date:	19th June 2007
Version:	V1.0
Document Name:	data-consultancy-report-final.doc
Notes:	

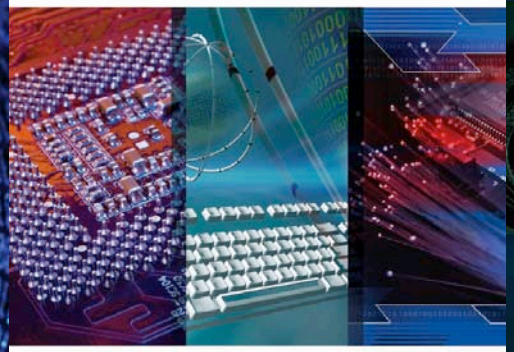
Harnessing the Power of Digital Data for Science and Society



Report of the Interagency Working Group on Digital Data to the Committee on Science of the National Science and Technology Council
January 2009

RCUK Review of e-Science 2009

BUILDING A UK FOUNDATION FOR THE TRANSFORMATIVE ENHANCEMENT OF RESEARCH AND INNOVATION



RESEARCH COUNCILS UK

THE ROYAL SOCIETY

Sustainable Economics for a Digital Planet:

Ensuring Long-Term Access to Digital Information

February 2010

Final Report of the Blue Ribbon Task Force on Sustainable Digital Preservation and Access

Data Sharing Policy – National Science Foundation

NSF Data Sharing Policy

- 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 VI.D.4.

NSF Data Management Plan Requirements

- Beginning January 18, 2011, proposals submitted to NSF 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.j for full policy implementation.

Data Sharing Policy – UK agencies

Wellcome Trust

- [It is] our expectation that all our funded researchers should maximize access to their research data with as few restrictions as possible. It requires applicants whose proposed research will generate data that hold significant value as a resource for the wider research community to submit a data management and sharing plan as part of the application process. Our policy aligns with those of other organizations - such as the MRC, BBSRC, the US National Institutes of Health and the NCRI Cancer Informatics initiative.

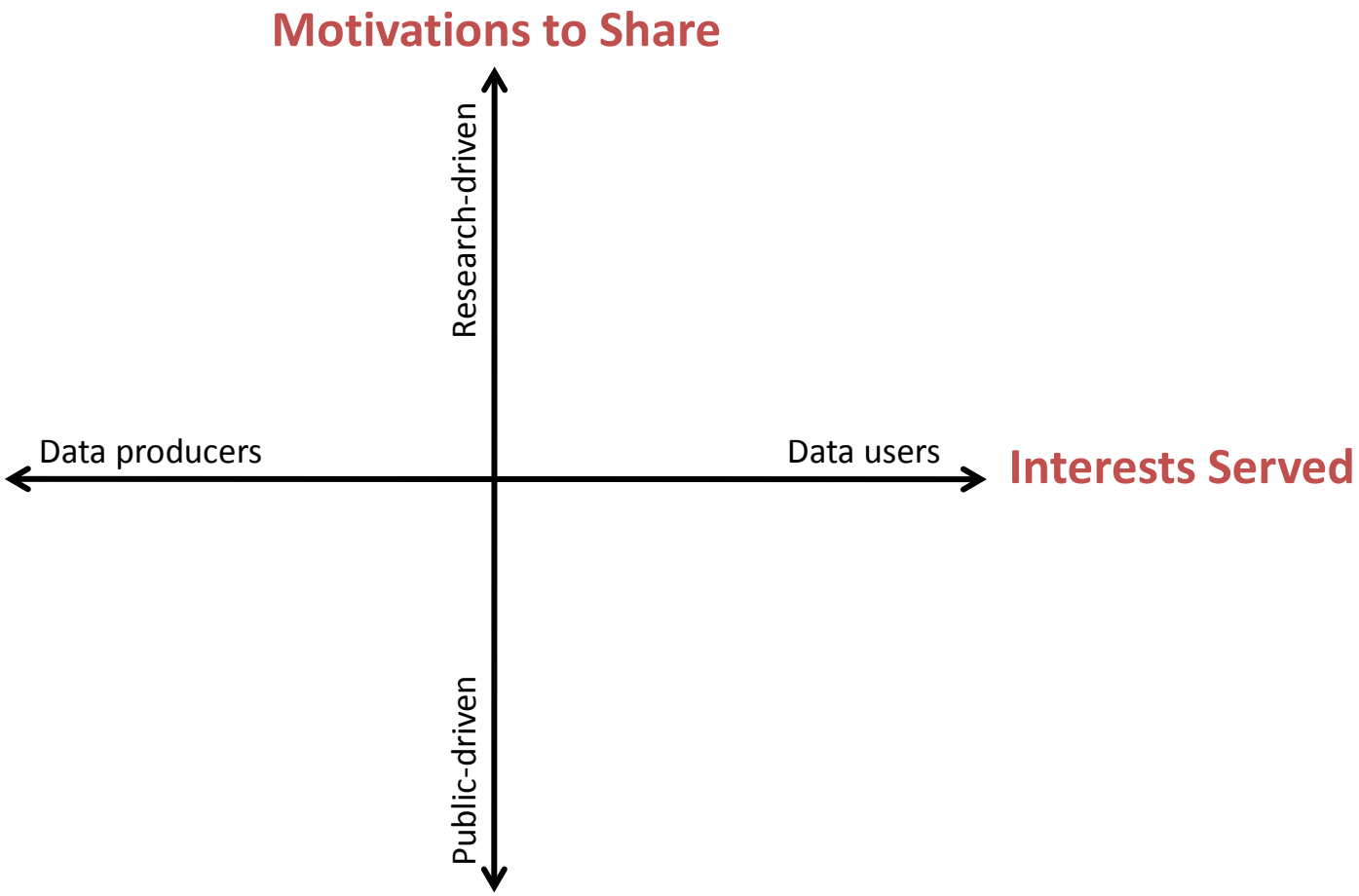
Economic and Social Research Council

- Our Research Data Policy is built upon the principle adopted by the OECD stating that publicly-funded research data are a public good, produced in the public interest and, therefore, should be openly available to a maximum extent possible. ESRC is, therefore, committed to long-term preservation, high quality data management and strengthening the provision for secondary data analysis.

Data


Experimental
Computational
Observational
Records
records
Birth
Maps
Data
Books
sensor
xrays
media
Opinion
indicators
readings
molecules
spreadsheets
lab
visualizations
dna
polls
Church
historical
events
artifacts
Field
types
water
Articles
marriage
temperature
experiments
Experiments
yearbooks
specimens
spectra
Census
images
structures
Voting
documents
time
letters
field
Diaries
genomes
objects
Surveys
Economic
laboratory
papers
Newspapers
Demographic
protein
Photographs
weather
photographs
ground
surveys
models
Court
interviews
Mass
notebooks
death

Rationales for Sharing Research Data



Borgman, C. L. (2011, forthcoming). The conundrum of sharing research data. Journal of the American Society for Information Science and Technology. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1869155 . Figure by Jillian C. Wallis

1. Reproduce or verify research

http://chemistry.curtin.edu.au/research/index.cfm

Benzoic Acid	% yield		IR Peaks (cm^{-1})		Solid (C) or Oil (O) Product	Mp (^{0}C)
	Gross	Recrystallization	N-H	C=O		
Sodium benzoate		2.58	3327	1638	White C	79-89
Sodium benzoate			3337	1640&1600	O	
Sodium benzoate			3326	1642&1601	O	
Sodium benzoate	37.8		3274	1640	O	
p-nitro	51.84	10.59	3423	1693	Yellow C	152-157
m-nitro	37.38	5.43	3334	1694	Green C	152-157
Benzoic acid		7.44	3293	1642	White C	152-154
m-bromo		47.4	3316	1702	Green paste	
p-bromo		14.53	3344	1638	Pink C	164-166
p-chloro		29.69	3340	1638	Yellow C	
m-chloro		74.53	3410	1637	tan paste	
o-chloro		17.31	3422	1654	Tan C	
3,5-dinitro		44.53	3297	1647	Tan C	139-141
p-hydroxy		3.751	3401	1643	yellow/green C	210
p-amino		8.475	3411	1645	Dark O	
o-methoxy		42.49	3412	1646	Yellow O	

http://serc.carleton.edu/cismi/broadaccess/groupwork.html

2. Public monies serve the public good

3. Others can ask new questions

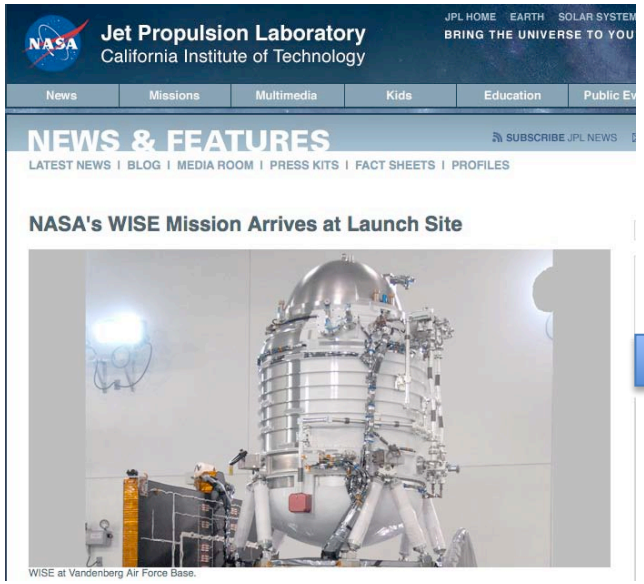
data

discovery

http://annualreport.ucdavis.edu/2008/images/photos/discovery.jpg

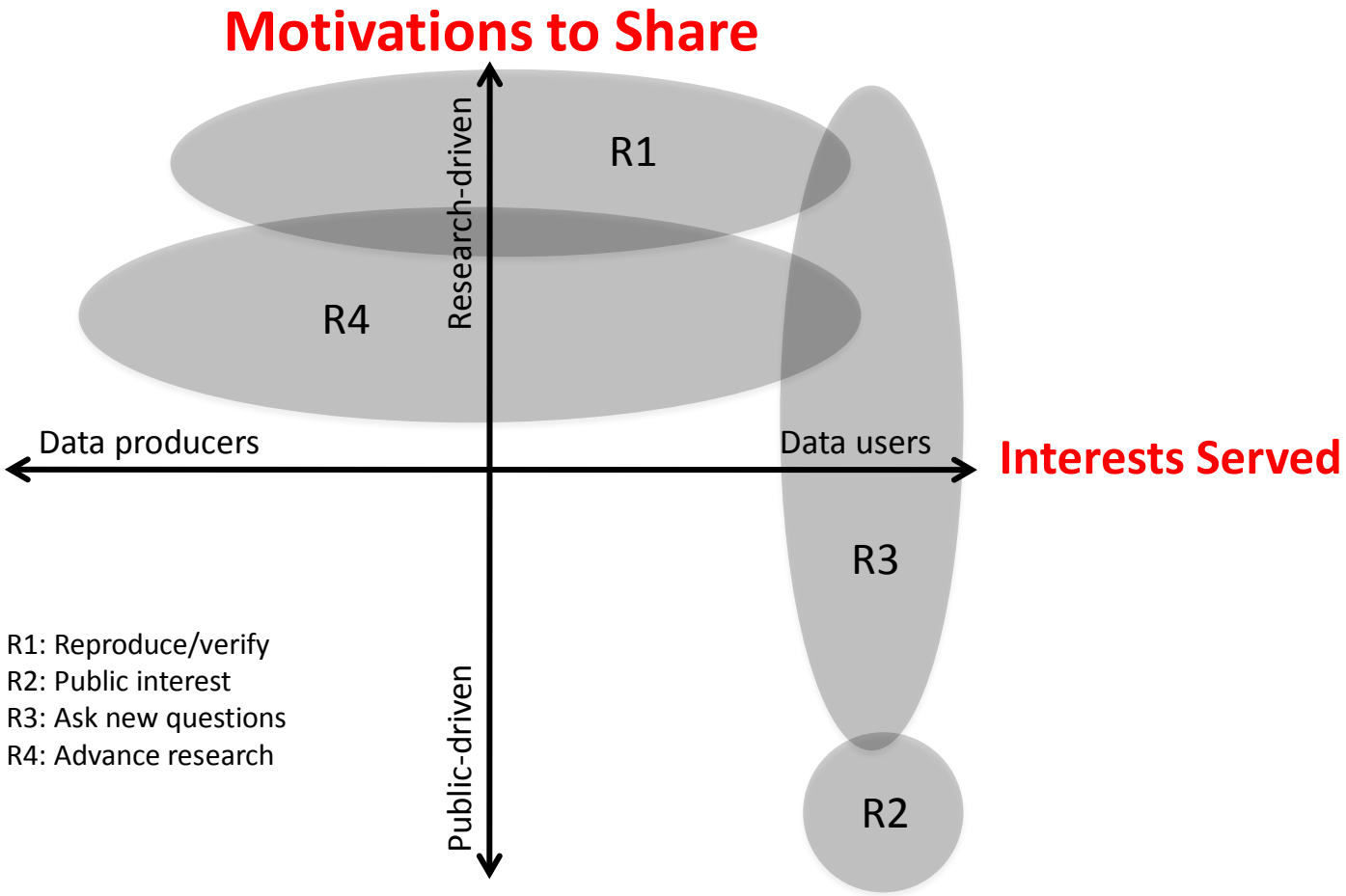
http://digitalassetmanagement.org.uk/2010/02/01/the-winds-of-change-are-blowing-in-the-clouds-favor/

4. Data curation advances innovation



International Virtual Observatory Alliance

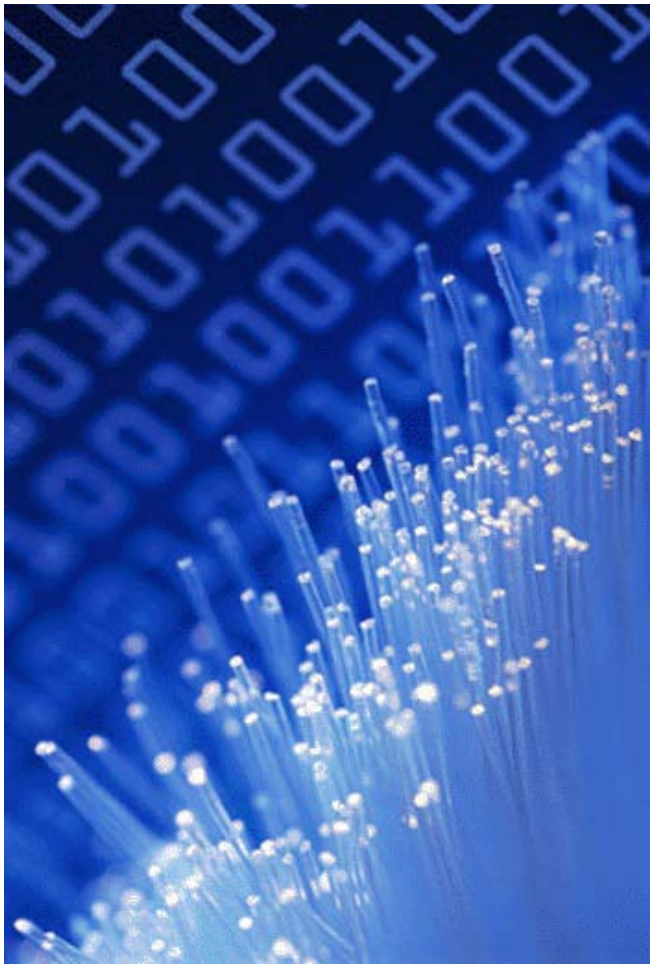
Rationales for Sharing Research Data



Borgman, C. L. (2011, forthcoming). The conundrum of sharing research data. Journal of the American Society for Information Science and Technology. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1869155 . Figure by Jillian C. Wallis

Infrastructure for research data

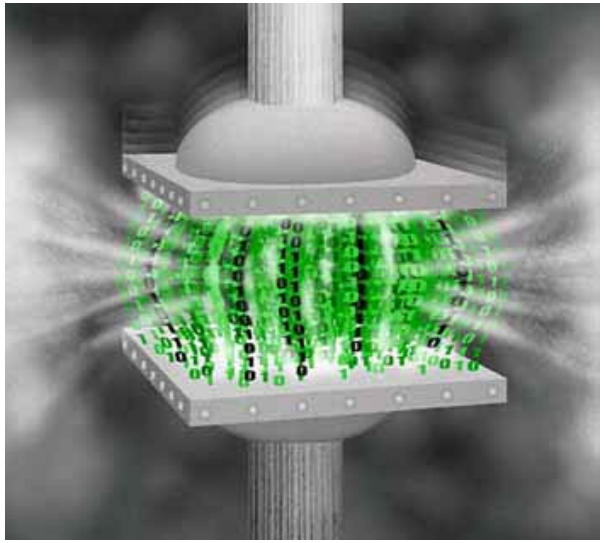
- Social practice
- Usability
- Identity
- Persistence
- Discoverability
- Provenance
- Relationships
- Intellectual property
- Policy



http://datalib.ed.ac.uk/GRAPHICS/blue_data.gif

Conclusions

- Research policy:
  - Data are intellectual property to be managed and exploited
- Funding agencies:
  - Research data are public assets
- Scholarly communication:
  - Research data are process and product
- Publishers:
  - Data are not publications
- Libraries:
  - Research data are the new special collections
- Researchers:
  - Some data will be shared, with some people, some of the time
- Internet research:
  - What are data, to whom, when, why, and to what ends?



http://plus.maths.org/content/text-bytes-and-videotape

Acknowledgements

- National Science Foundation
  - CENS: Cooperative Agreement #CCR-0120778, D.L. Estrin, UCLA, PI.
  - CENS Education Infrastructure: #ESI- 0352572, W.A. Sandoval, PI; C.L. Borgman, co-PI.
  - Towards a Virtual Organization for Data Cyberinfrastructure, #OCI-0750529, C.L. Borgman, UCLA, PI; G. Bowker, Santa Clara University, Co-PI; T. Finholt, University of Michigan, Co-PI.
  - Monitoring, Modeling & Memory: Dynamics of Data and Knowledge in Scientific Cyberinfrastructures: #0827322, P.N. Edwards, UM, PI; Co-PIs C.L. Borgman, UCLA; G. Bowker, SCU; T. Finholt, UM; S. Jackson, UM; D. Ribes, Georgetown; S.L. Star, SCU)
  - Data Conservancy: OCI0830976, Sayeed Choudhury, PI, Johns Hopkins University.
- Microsoft External Research: Tony Hey, Lee Dirks, Catherine van Ingen, Catherine Marshall
- Conundrum paper comments: CENS Data Practices team at UCLA – David Fearon, Matthew Mayernik, Katie Shilton, Jillian Wallis, and Laura Wynholds; Paul Uhlir of the National Academies.

Microsoft