

# UCLA

## Presentations

### Title

Big Data, Little Data, noData: Scholarship in the Networked World

### Permalink

<https://escholarship.org/uc/item/6vt1h4wt>

### Author

Borgman, Christine L.

### Publication Date

2014-02-04

### Copyright Information

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

# Big Data, Little Data, No Data: Scholarship in the Networked World

---

Christine L. Borgman

Professor and Presidential Chair in Information Studies  
University of California, Los Angeles

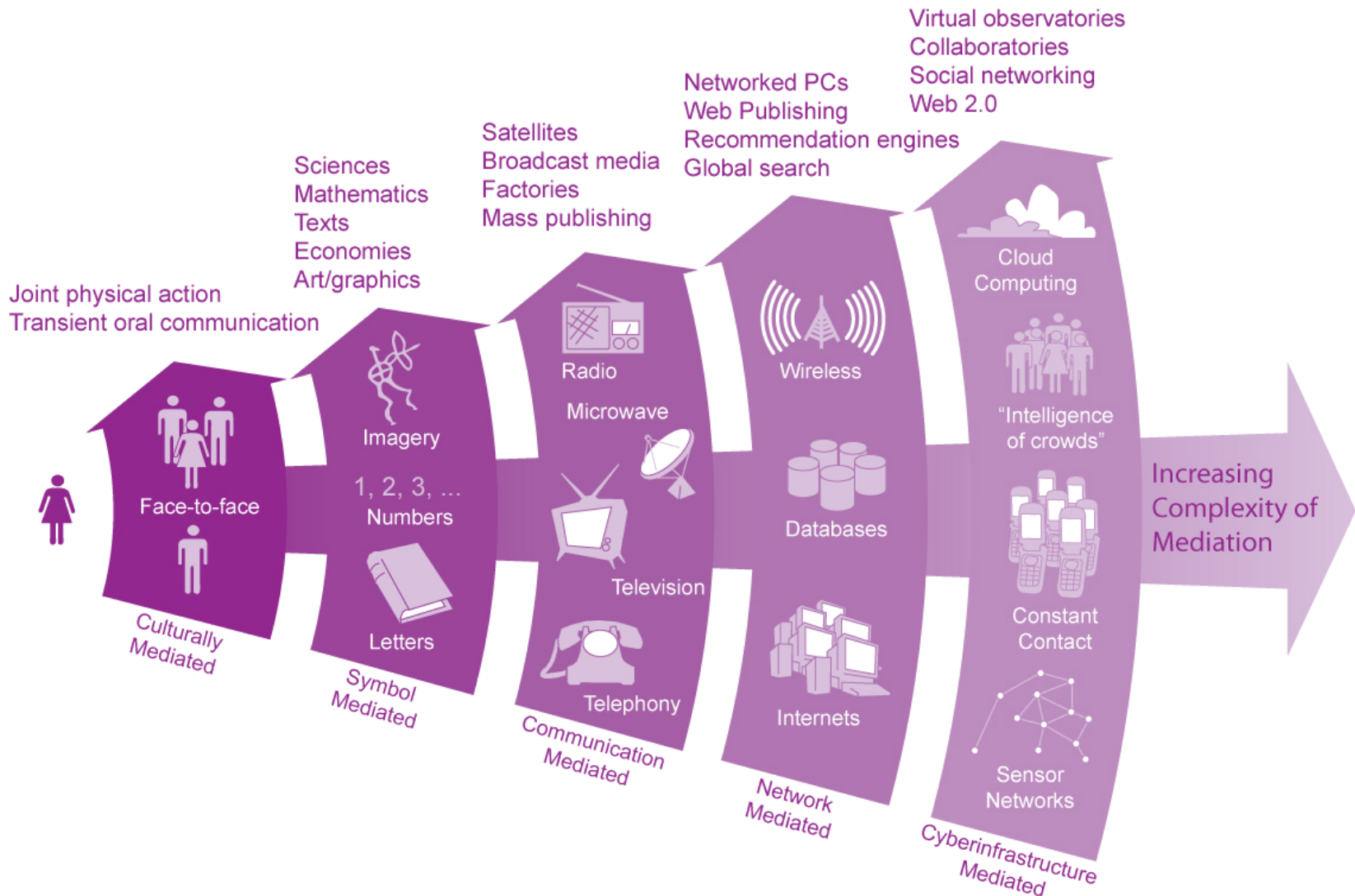


3-6 February 2014

Melbourne Convention & Exhibition Centre

streaming with possibilities

# Technological advances in mediated communication



Borgman, C. L., Abelson, H., Dirks, L., Johnson, R., Koedinger, K. R., Linn, M. C., ... Szalay, A. (2008). *Fostering Learning in the Networked World: The Cyberlearning Opportunity and Challenge*. National Science Foundation. [http://www.nsf.gov/publications/pub\\_summ.jsp?ods\\_key=nsf08204](http://www.nsf.gov/publications/pub_summ.jsp?ods_key=nsf08204)

# Science

12 September 2008 | \$10



AAAS

# Science

11 February 2011 | \$10



# Neelie Kroes, VP European Commission:



To collect, curate, preserve and make available ever-increasing amounts of scientific data, new types of infrastructures will be needed. The potential benefits are enormous but the same is true for the costs. We therefore need to lay the right foundations and the sooner we start the better.

Wood, J., Andersson, T., Bachem, A., Best, C., Genova, F., Lopez, D. R., ... Hudson, R. L. (2010). *Riding the wave: How Europe can gain from the rising tide of scientific data*. Final report of the High Level Expert Group on Scientific Data. Retrieved from <http://cordis.europa.eu/fp7/ict/e-infrastructure/docs/hlg-sdi-report.pdf>

**ANDS Home**

**About ANDS**

[Our Approach](#)

[Better Data: Better Research](#)

[Our Governance](#)

[Our Publications](#)

**Partners & Communities**

**Data Management**

**Metadata**

**Discovery, Access, Reuse**

**Technical Resources**

**Guides, Training, Support**

**ANDS Services**

**News & Events**

## Better Data: Better Research

### Why manage data?

- Preserve the integrity of the research
- Allow data to be made available for others to use
- Assist researchers to reduce the risk of data loss
- Secure continued access to the value in data

### Why connect data?

- Interlink data to people to projects to publications
- Improve the discoverability of data
- Tie data to research achievements
- Provide richer context for data value

### Why make data discoverable?

- Enable the demonstration of research excellence
- Allow researchers to build upon existing data, instead of recreating it
- Foster innovation
- Provide the ability to solve big problems across discipline boundaries

### Why reuse data?

- Verification of research claims
- New discoveries from existing data
- Integration of sets of data for new analysis
- Re-analysis of expensive, rare or unrepeatable investigations
- Reduction of duplicated effort



# Open access policies

- Australian Research Council
  - Code for the Responsible Conduct of Research
  - Data management plans
- National Science Foundation
  - Data sharing requirements
  - Data management plans
- U.S. Federal policy-2013
  - Open access to publications
  - Open access to data
- European Union
  - European Open Data Challenge
  - OpenAIRE
- Research Councils of the UK
  - Open access publishing
  - Provisions for access to data



National Science Foundation  
WHERE DISCOVERIES BEGIN



Australian Government

National Health and Medical Research Council

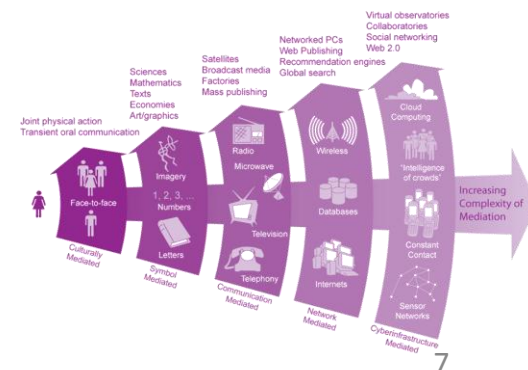
Supported by  
**welcome**trust

Policy RECommendations for Open Access to Research Data in Europe



# Big Data, Little Data, No Data: Scholarship in the Networked World\*

- Section I: Data and Scholarship
  - Ch 1: Provocations
  - Ch 2: What are Data?
  - Ch 3: Data Scholarship
  - Ch 4: Data Diversity
- Section II: Case Studies in Data Scholarship
  - Ch 5: Data Scholarship in the Sciences
  - Ch 6: Data Scholarship in the Social Sciences
  - Ch 7: Data Scholarship in the Humanities
- Section III Data Policy and Practice
  - Ch 8: Releasing, Sharing, and Reusing Data
  - Ch 9: Credit, Attribution, and Discovery
  - Ch 10: What to Keep and Why



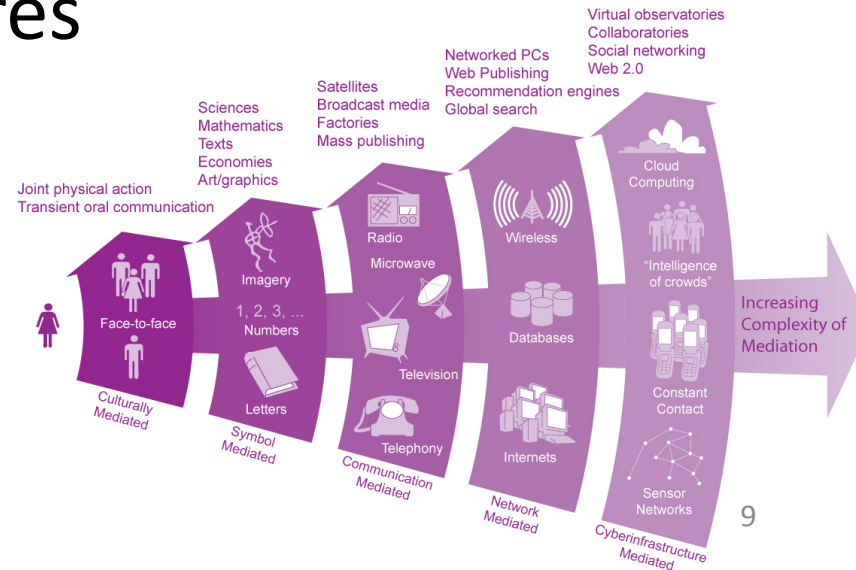


# Provocations

1. How do rights, responsibilities, and risks associated with research data vary by discipline and by stakeholder?
2. How can data be exchanged across domains, contexts, and over time?
3. How do publications and data differ?
4. What are scholars' motivations to share their data?
5. What kinds of expertise are required to manage, curate, and sustain access to research data?
6. How can knowledge infrastructures adapt to the needs of generations of scholars and the demands of competing stakeholders?

# Libraries, librarians, and data

1. Data are not publications
2. Data are not natural objects
3. Data are representations
4. Data sharing and reuse depend on knowledge infrastructures



PHILOSOPHICAL  
TRANSACTIONS:  
GIVING SOME  
ACCOMPT  
OF THE PRESENT  
Undertakings, Studies, and Labours  
OF THE  
INGENIOUS  
IN MANY  
CONSIDERABLE PARTS  
OF THE  
WORLD.

---

*Vol. I.*

For *Annos* 1665, and 1666.

---

In the SAVOY,  
Printed by T. N. for John Martyn at the Bell, a little with-  
out Temple-Bar, and James Alsby in South-Lane!  
Printed to the Royal Society.

*Vol. I. p. 377.  
U. 2. p. 332. p. 332  
U. 3. p. 602. p. 332*



Brick inscribed with the Sutra on Dependent Origination *Gorakhpur district, late 5th century - early 6th century AD. Ashmolean Museum*

# Open access publishing

PHILOSOPHICAL  
TRANSACTION  
GIVING SOME  
ACCOMP  
OF THE PRESENT  
Undertakings, Studies, and Labou  
OF THE  
INGENIOU



[Subject Areas](#)

[For Authors](#)

[About Us](#)

Search



[advanced search](#)

OPEN ACCESS PEER-REVIEWED

RESEARCH ARTICLE

9,703

VIEWS

2

CITATIONS

12

SAVES

251

SHARES

## If We Share Data, Will Anyone Use Them? Data Sharing and Reuse in the Long Tail of Science and Technology

Jillian C. Wallis, Elizabeth Rolando, Christine L. Borgman

Published: July 23, 2013 • DOI: 10.1371/journal.pone.0067332



[Home](#)

[About](#)

[Media](#)

[Get Involved](#)

[Committees](#)

Metrics

Comments

Related Content

Download PDF

Print

Share

Subject Areas

Data management

Data processing

Oceans

Research laboratories

Researchers

Science policy

Scientists

Seismology

### Editorial Recruitment

The OLH is now recruiting editors for each of our disciplinary fields. If you have academic editorial expertise and would like to get involved in open access publishing, please get in ...

### Latest News



April 17, 2013 | 0 Comments and 0 Reactions

### BIS Select Committee on Open Access

On 16th April, OLH Project Director Dr Martin Eve was invited to give evidence at the Business, Innovation and Skills Committee, which is appointed by the UK House of Commons to examine the administration, expenditure and policy of the Department for Business, Innovation and Skills (BIS) and its associated public bodies, including Ofcom and the

### ROADMAP FOR TECHNICAL PILOT

### EDITORIAL RECRUITMENT

### FORMS OF INNOVATION

Search...

### WELCOME

Welcome to the Open Library of Humanities (OLH). This site aims to give the background to, and rationale for, our vision of building a low cost, sustainable, Open Access future for the humanities. Please feel free to look around the site and get in touch if you'd like to be involved.

### RECENT POSTS

[Roadmap for Technical Pilot](#)

practices to share and reuse data will inform the design of infrastructure to action, management, and discovery in the long tail of science and technology. Tech domains in which data tend to be local in character, minimally structured, documented. We report on a ten-year study of the Center for Embedded Networked Computing, a National Science Foundation Science and Technology Center. We found researchers are willing to share their data, but few are asked to do so, and in only a few cases do their funders or journals require them to deposit data. Few repositories exist in CENS research areas.. Data sharing tends to occur only through informal channels. CENS researchers obtain data from repositories, and occasionally from individuals, to provide context, calibration, or other forms of background for their CENS researchers nor those who request access to CENS data appear to be willing to share data if they receive credit and retain first rights to publish their research. The practices of releasing, sharing, and reusing of data in CENS reaffirm the gift culture of high goods are bartered between trusted colleagues rather than treated as

# Open dissemination

Cornell University Library

We gratefully acknowledge support from the Simons Foundation and member institutions

arXiv.org

Search or Article-id  [\(help | Advanced search\)](#)

[All papers](#) [Go](#)

[Login](#)

Open access to 846,777 e-prints in Physics, Mathematics, Computer Science, Quantitative Biology, Quantitative Finance and Statistics

Subject search and browse: [Physics](#) [Search](#) [Form Interface](#) [Catchup](#)

29 Aug 2012: Simons Foundation funds new arXiv sustainability model  
See cumulative "What's New" pages. Read robots beware before attempting any automated download

## Physics

- **Astrophysics (astro-ph new, recent, find)**  
includes: Cosmology and Extragalactic Astrophysics; Earth and Planetary Astrophysics; Solar and Stellar Astrophysics
- **Condensed Matter (cond-mat new, recent, find)**  
includes: Disordered Systems and Neural Networks; Materials Science; Mesoscale & Mechanics; Strongly Correlated Electrons; Superconductivity
- **General Relativity and Quantum Cosmology (gr-qc new, recent, find)**
- **High Energy Physics - Experiment (hep-ex new, recent, find)**
- **High Energy Physics - Lattice (hep-lat new, recent, find)**
- **High Energy Physics - Phenomenology (hep-ph new, recent, find)**
- **High Energy Physics - Theory (hep-th new, recent, find)**
- **Mathematical Physics (math-ph new, recent, find)**
- **Nonlinear Sciences (nlin new, recent, find)**  
includes: Adaptation and Self-Organizing Systems; Cellular Automata and Lattice &
- **Nuclear Experiment (nucl-ex new, recent, find)**
- **Nuclear Theory (nucl-th new, recent, find)**
- **Physics (physics new, recent, find)**

**SSRN SOCIAL SCIENCE RESEARCH NETWORK**  
Tomorrow's Research Today

SSRN USER HQ

Find me a user? Free Registration

USER ID  [Go](#)

PASSWORD  [Go](#)

Forgot ID or Password? Help

Accounting ARN Research Network

Anthropology & Archaeology Research Network

Cognitive Science Network

Corporate Governance Network

Economics ERN Research Network

Entrepreneurship Research & Policy Network

Financial Economics Network

Health Economics Network

Information Systems & eBusiness Network

Innovation Research & Policy Network

Leadership Research Network

Legal Scholarship Network

Management Research Network

Marketing Research Network

Negotiations Research Network

Home Search Browse Submit Subscribe Shopping Cart My Briefcase

**SSRN eLibrary**

co-hosted by

CHICAGO BOOTH european corporate governance institute

KOREA UNIVERSITY Stanford Law School

Top Papers Top Authors Top Institutions

Search Browse

Research Paper Series

Partners in Publishing

Institution Home Pages

**Leading Social Science Research Delivered Daily**

Social Science Research Network (SSRN) is devoted to the rapid worldwide dissemination of social science research and is composed of a number of specialized research networks in each of the social sciences. We have received several excellence awards for our web site.

Each of SSRN's networks encourages the early distribution of research results by publishing Submitted abstracts and by soliciting abstracts of top quality research papers around the world. We now have hundreds of journals, publishers, and institutions in Partners in Publishing that provide working papers for distribution through SSRN's eLibrary and abstracts

**SSRN's Objective and Commitments to Users**

20th Anniversary Message from: Michael C. Jensen, SSRN Chairman

**Recent Announcements**

SSRN has again been named the Number 1 Open Access Repository in the World (for July, 2012) by the Ranking Web of World Repositories. Our thanks to all of the SSRN community who helped make this happen.

UNDP Human Development Report Office Joins ERN Government & Public Agency Research Paper Series

Announcing New Cyberspace Law - Student Authors and Intellectual Property Law - Student Authors eJournals on SSRN

Announcing IICM 2012 11th Indian Institute of Capital Markets Conference Online Proceedings on SSRN

Regulatory Institutions Network (RegNet) Joins Law Research Centers Papers

Capitalism & Society Published in FEN Partners in Publishing Journals

Singapore Management University Joins Accounting Research Centers Papers

Search Contribute Browse Help

**ORA**  
Oxford University Research Archive

Research publications from the University of Oxford

**SEARCH** **CONTRIBUTE** **BROWSE**

[Metadata Only](#) [SEARCH](#) [Help](#)

**Top Downloads**

Malcolm David Birdling, (2012). Correction of miscarriages of justice in New Zealand and England. DPhil. University of Oxford.

John Baines, The Sebekholpe VIII inundation stela: an additional fragment. Acta Orientalia, 37, 11-20.

John Baines, The inundation stela of Sebekholpe VIII. Acta Orientalia, 36, 39-54.

Thomas Outdridge, (2011). Coarse-grained modelling of DNA and DNA self-assembly. DPhil. University of Oxford.

Elna Klipi, (2010). The education of children of immigrants in Finland. DPhil. University of Oxford.

[Follow @oxforduni\\_repo](#)

**Latest Additions**

Nicole Gilroy and Andrew Honey, (2012). The conservation of two composite Anselm manuscripts from the twelfth century: two contrasting approaches?.

Debopam Bhattacharya, Shin Kanaya and Margaret Stevens, (2013). Are university admissions academically fair?.

Mark Irwin, (2013). The synthesis and characterisation of metal complexes containing chemically reduced bipyridyl ligand systems. DPhil. University of Oxford.

Andrea Polo, (2012). Essays in corporate restructuring, reputation and law. DPhil. University of Oxford.

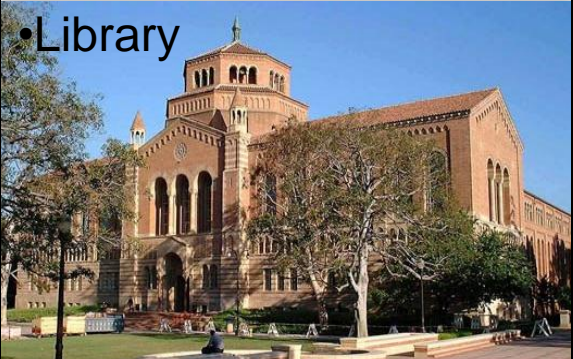
Keith Hyams and Tina Fawcett, (2013-March/April). The ethics of carbon offsetting. WIREs Climate Change, 4 (2), 91-98.

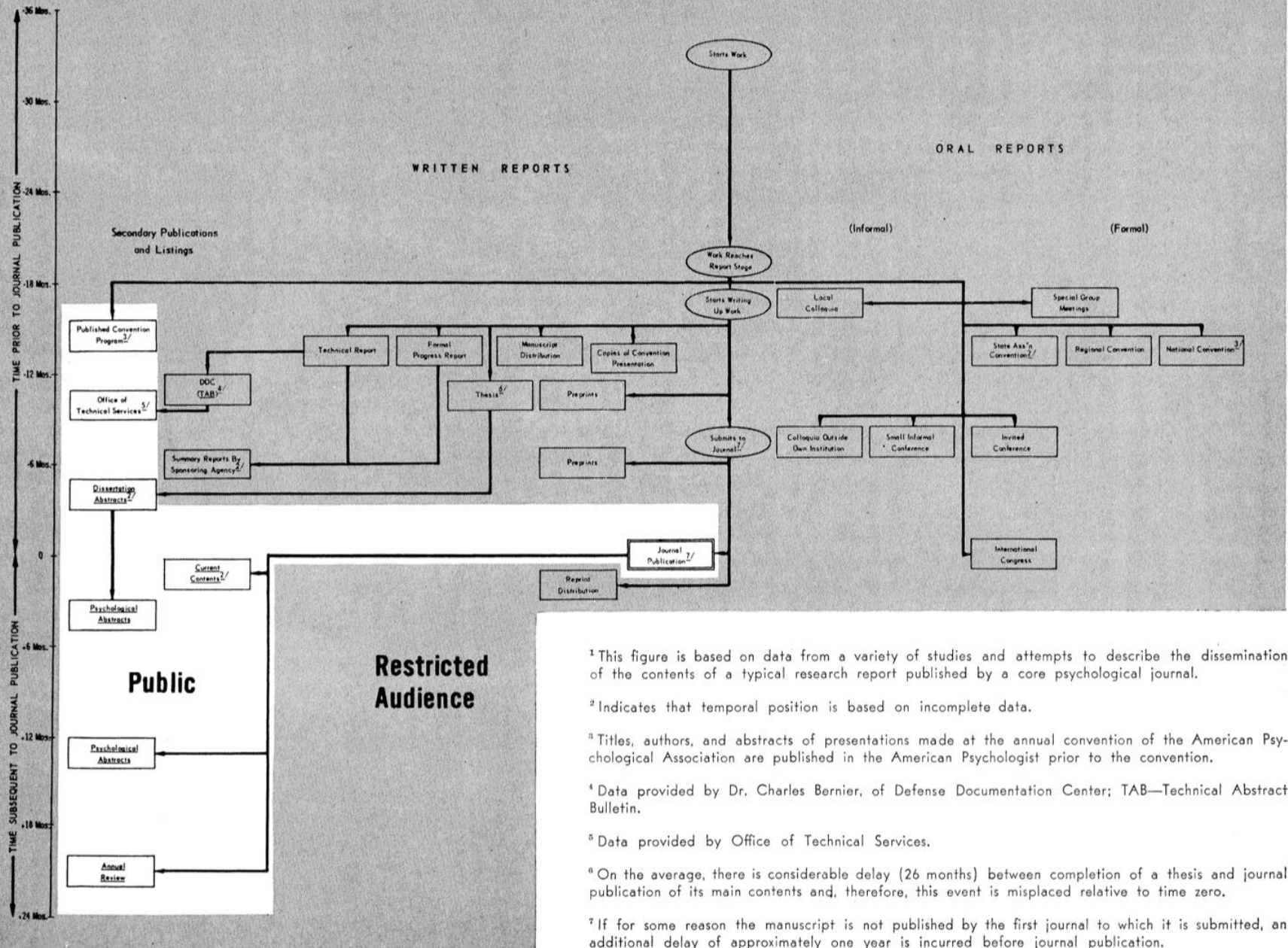
Site Copyright © 2008 Oxford University Research Archive. All rights reserved.

Site not yet optimized for viewing in mobile devices.

Downloads

# Publications in Scholarly Communication

Function	Print	Digital
<b>Legitimization</b> Authority, quality, priority, trustworthiness	<ul style="list-style-type: none"> <li>•Peer review</li> </ul>	<ul style="list-style-type: none"> <li>•Peer review</li> </ul>
<b>Dissemination</b> Awareness, diffusion, publicity	<ul style="list-style-type: none"> <li>•Publisher</li> <li>•Pre-print distribution                             <ul style="list-style-type: none"> <li>•Copy</li> <li>•Mail</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>•Publisher</li> <li>•Pre-print distribution                             <ul style="list-style-type: none"> <li>•Post on Web</li> <li>•Deposit</li> </ul> </li> </ul>
<b>Access, preservation, curation</b> Availability, discovery, retrieval, persistence	<ul style="list-style-type: none"> <li>•Library</li> </ul> 	<ul style="list-style-type: none"> <li>•Library</li> <li>•Publisher</li> <li>•Repository</li> <li>•Homepage</li> </ul>

THE DISSEMINATION OF SCIENTIFIC INFORMATION IN PSYCHOLOGY<sup>v</sup>

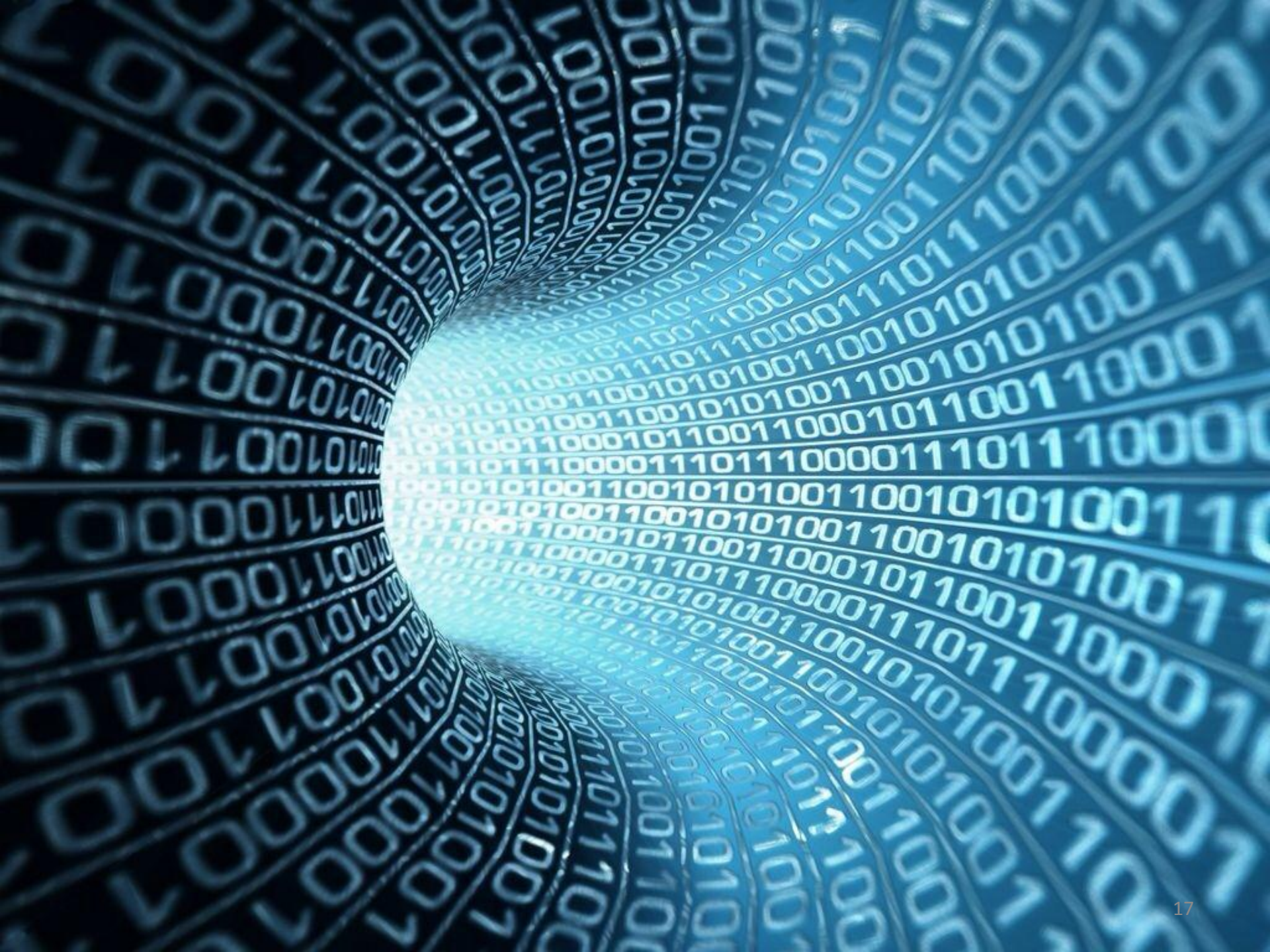


# Open access to scholarly literature

- Open access literature is digital, online, free of charge, and free of most copyright and licensing restrictions.\*
- Principles
  - Scholarly authors hold the copyright on their work
  - Scholarly authors write for impact, not revenue

\*Suber, P. (2012). *Open Access*. Cambridge MA: MIT Press.

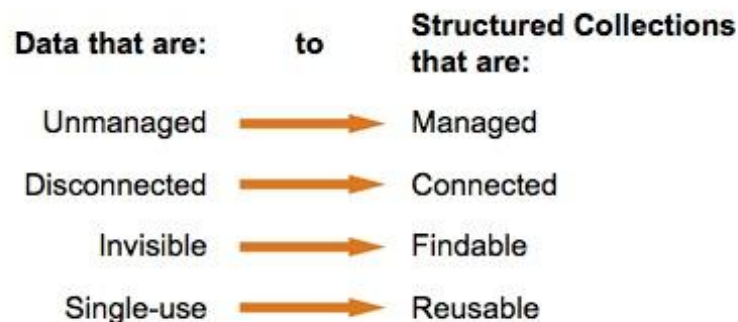
<http://mitpress.mit.edu/books/open-access>



# Australian National Data Service

***Our Vision: More Australian researchers reusing research data more often***

**ANDS is enabling the transformation of:**



[More>>](#)

## **Australian Research Data Commons**

ANDS is building the **Australian Research Data Commons**: a cohesive collection of research resources from all research institutions, to make better use of Australia's research data outputs.

## **Research Data Australia**

**Research Data Australia**, ANDS' flagship service, provides a comprehensive window into the Australian Research Data Commons. It is an Internet-based discovery service designed to provide rich connections between data, projects, researchers and institutions, and promote visibility of Australian research data collections in search engines.

Search for research data on [researchdata.ands.org.au](https://researchdata.ands.org.au)

# Open Data

- A piece of data or content is open if anyone is free to use, reuse, and redistribute it — subject only, at most, to the requirement to attribute and/or share-alike (Open Data Commons, 2013).
- Data that meets the criteria of intelligent openness. Data must be accessible, useable, assessable and intelligible. (Royal Society, 2012)
- Openness, flexibility, transparency, legal conformity, protection of intellectual property, formal responsibility, professionalism, interoperability, quality, security, efficiency, accountability, and sustainability. *OECD Principles and Guidelines for Access to Research Data from Public Funding* (2007)

# OA publications vs. data

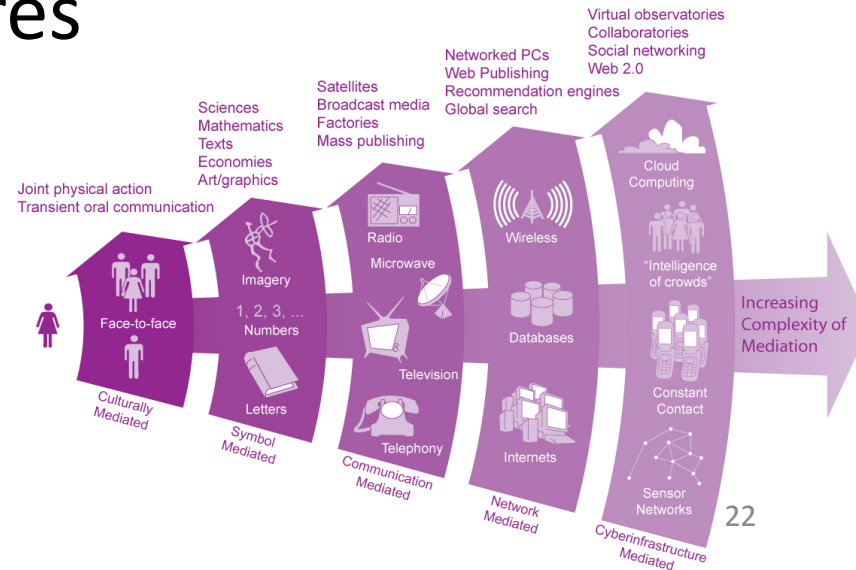
- Open access to literature principles (Suber, 2012)
  - Scholarly authors hold the copyright on their work
  - Scholarly authors write for impact, not revenue
- Open access to data
  - Rights in data are unclear and often contested
  - Papers are arguments for which data are evidence

# Data in Scholarly Communication

Function	Reported in a publication	Contributed to a data repository
<b>Legitimization</b> Authority, quality, priority, trustworthiness	<ul style="list-style-type: none"> <li>•Peer review in context                             <ul style="list-style-type: none"> <li>•Quality of method</li> <li>•Evidence for conclusions</li> <li>•Verify, reanalyze?</li> </ul> </li> <li>•Author reputation</li> </ul>	<ul style="list-style-type: none"> <li>•Peer review                             <ul style="list-style-type: none"> <li>•Quality of metadata, documentation</li> <li>•“test drive” the data</li> </ul> </li> </ul>
<b>Dissemination</b> Awareness, diffusion, publicity	<ul style="list-style-type: none"> <li>•Description in a publication</li> </ul>	<ul style="list-style-type: none"> <li>•Repository publisher</li> </ul>
<b>Access, preservation, curation</b> Availability, discovery, retrieval, persistence	<ul style="list-style-type: none"> <li>•Request to author</li> <li>•Author maintains own data</li> <li>•Author directs requestor to data source</li> </ul>	<ul style="list-style-type: none"> <li>•Repository                             <ul style="list-style-type: none"> <li>•Immediate access</li> <li>•Embargo period</li> </ul> </li> <li>•Curation responsibility</li> </ul>

# Libraries, librarians, and data

1. Data are not publications
2. Data are not natural objects
3. Data are representations
4. Data sharing and reuse depend on knowledge infrastructures



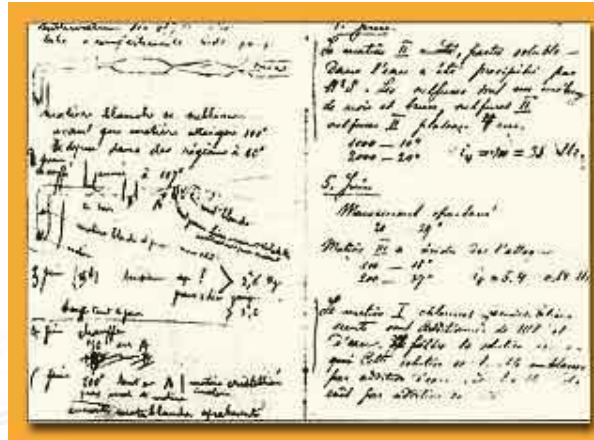
# What are data?



NASA Astronomy Picture of the Day

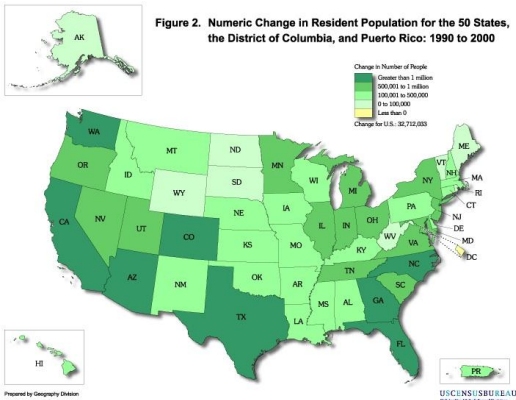


hudsonalpha.org



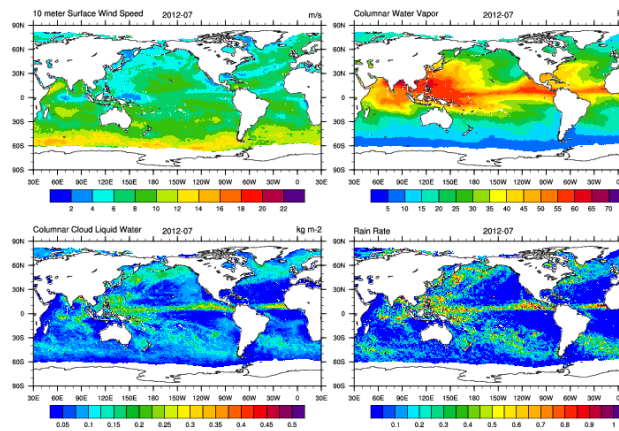
Marie Curie's notebook aip.org

Figure 2. Numeric Change in Resident Population for the 50 States, the District of Columbia, and Puerto Rico: 1990 to 2000



<http://www.census.gov/population/cen2000/map02.gif>

Monthly Mean: f17\_ssmis\_201207v7.nc



ncl.ucar.edu

Date: 1/2.07.75 Place: Sakaltutan Zafor

He will grow old in his present house; new house is for sons - 5 sons. Not sure they want to live in village. He will only build another if they want him to. eS came from Germany and did the plastering. He arranged the carpentry in Kayseri. Çok para gitti. (much money went) Has a tractor.

Date: July 1980 Place: Sakaltutan Zafor:

Household now Zafor and wife; Nazif Unal and wife and youngest son, still a boy. They run two dolmuş; one with a driver from Süleymanlı. Goes in and out once a day. He gets 8,000 a month. Zafor then said, keskin deOil. (not sharp - i.e.? not profitable) I said he did very well on 8,000 TL with only two journeys a day. Nazif Unal has "bought" a Durak (dolmuş stop) from Belediye and works all day in Kayseri.

[http://onlineqda.hud.ac.uk/Intro\\_QDA/Examples\\_of\\_Qualitative\\_Data.php](http://onlineqda.hud.ac.uk/Intro_QDA/Examples_of_Qualitative_Data.php)





## LETTERS

## A role for self-gravity at multiple length scales in the process of star formation

Alyssa A. Goodman<sup>1,2</sup>, Erik W. Rosolowsky<sup>3,5</sup>, Michelle A. Borkin<sup>1,†</sup>, Jonathan B. Foster<sup>2</sup>, Michael Halle<sup>1,4</sup>, Jens Kauffmann<sup>1,2</sup> & Jaime E. Pineda<sup>2</sup>

Self-gravity plays a decisive role in the final stages of star formation, where dense cores (size  $\sim 0.1$  parsecs) inside molecular clouds collapse to form star-plus-disk systems<sup>1</sup>. But self-gravity's role at earlier times (and on larger length scales, such as  $\sim 1$  parsec) is unclear; some molecular cloud simulations that do not include self-gravity suggest that 'turbulent fragmentation' alone is sufficient to create a mass distribution of dense cores that resembles, and sets, the stellar initial mass function<sup>2</sup>. Here we report a 'dendrogram' (hierarchical tree-diagram) analysis that reveals that self-gravity plays a significant role over the full range of possible scales traced by <sup>13</sup>CO observations in the L1448 molecular cloud, but not everywhere in the observed region. In particular, more than 90 per cent of the compact 'pre-stellar cores' traced by peaks of dust emission are projected on the sky within one of the dendrogram's self-gravitating 'leaves'. As these peaks mark the locations of already-forming stars, or of those probably about to form, a self-gravitating cocoon seems a critical condition for their existence. Turbulent fragmentation simulations without self-gravity—even of unmagnetized isothermal material—can yield mass and velocity power spectra very similar to what is observed in clouds like L1448. But a dendrogram of such a simulation<sup>3</sup> shows that nearly all the gas in it (much more than in the observations) appears to be self-gravitating. A potentially significant role for gravity in 'non-self-gravitating' simulations suggests inconsistency in simulation assumptions and output, and that it is necessary to include self-gravity in any realistic simulation of the star-formation process on subparsec scales.

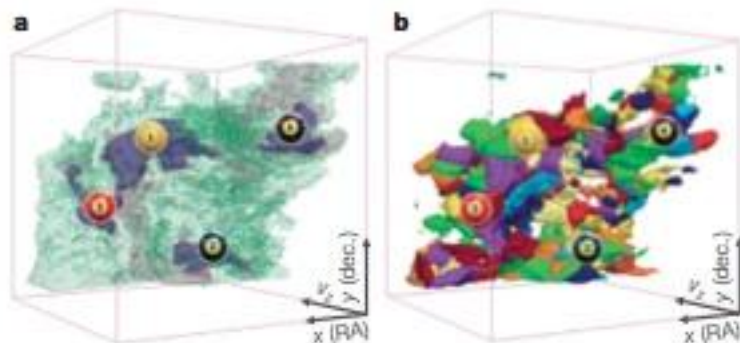
Spectral-line mapping shows whole molecular clouds (typically tens to hundreds of parsecs across, and surrounded by atomic gas) to be marginally self-gravitating<sup>4</sup>. When attempts are made to further break down clouds into pieces using 'segmentation' routines, some self-gravitating structures are always found on whatever scale is sampled<sup>5,6</sup>. But no observational study to date has successfully used one spectral-line data cube to study how the role of self-gravity varies as a function of scale and conditions, within an individual region.

Most past structure identification in molecular clouds has been explicitly non-hierarchical, which makes difficult the quantification of physical conditions on multiple scales using a single data set. Consider, for example, the often-used algorithm CLUMPFIND<sup>7</sup>. In three-dimensional (3D) spectral-line data cubes, CLUMPFIND operates as a watershed segmentation algorithm, identifying local maxima in the position-position-velocity (p-p-v) cube and assigning nearby emission to each local maximum. Figure 1 gives a two-dimensional (2D) view of L1448, our sample star-forming region, and Fig. 2 includes a CLUMPFIND decomposition of it based on <sup>13</sup>CO observations. As with any algorithm that does not offer hierarchically nested or

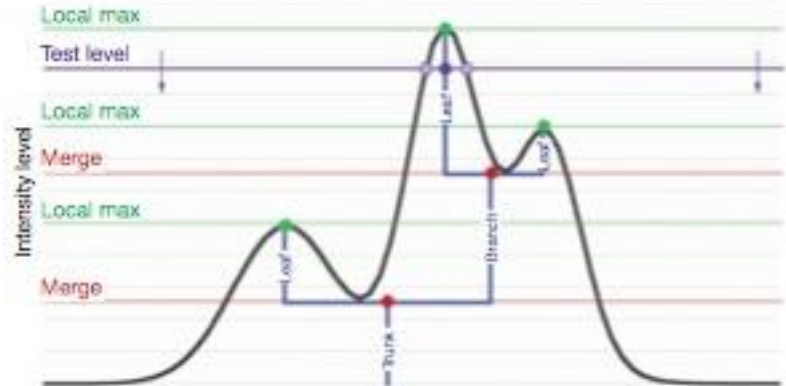
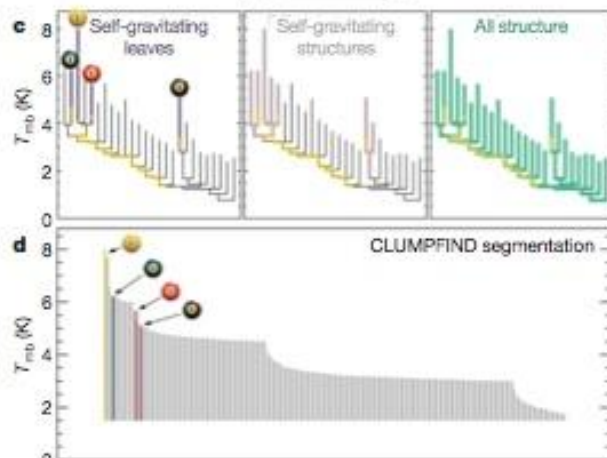
overlapping features as an option, significant emission found between prominent clumps is typically either appended to the nearest clump or turned into a small, usually 'pathological', feature needed to encompass all the emission being modelled. When applied to molecular-line



**Figure 1** | Near-infrared image of the L1448 star-forming region with contours of molecular emission overlaid. The channels of the colour image correspond to the near-infrared bands J (blue), H (green) and K (red), and the contours of integrated intensity are from <sup>13</sup>CO(1-0) emission<sup>8</sup>. Integrated intensity is monotonically, but not quite linearly (see Supplementary Information), related to column density<sup>9</sup>, and it gives a view of 'all' of the molecular gas along lines of sight, regardless of distance or velocity. The region within the yellow box immediately surrounding the protostar has been imaged more deeply in the near-infrared (using Calar Alto) than the remainder of the box (2MASS data only), revealing protostars as well as the scattered starlight known as 'Cloisidine'<sup>10</sup> and outflow (which appear orange in this colour scheme). The four billiard-ball labels indicate regions containing self-gravitating dense gas, as identified by the dendrogram analysis, and the leaves they identify are best shown in Fig. 2a. Asterisks show the locations of the four most prominent embedded young stars or compact stellar systems in the region (see Supplementary Table 1), and yellow circles show the millimetre-dust emission peaks identified as star-forming or 'pre-stellar' cores.



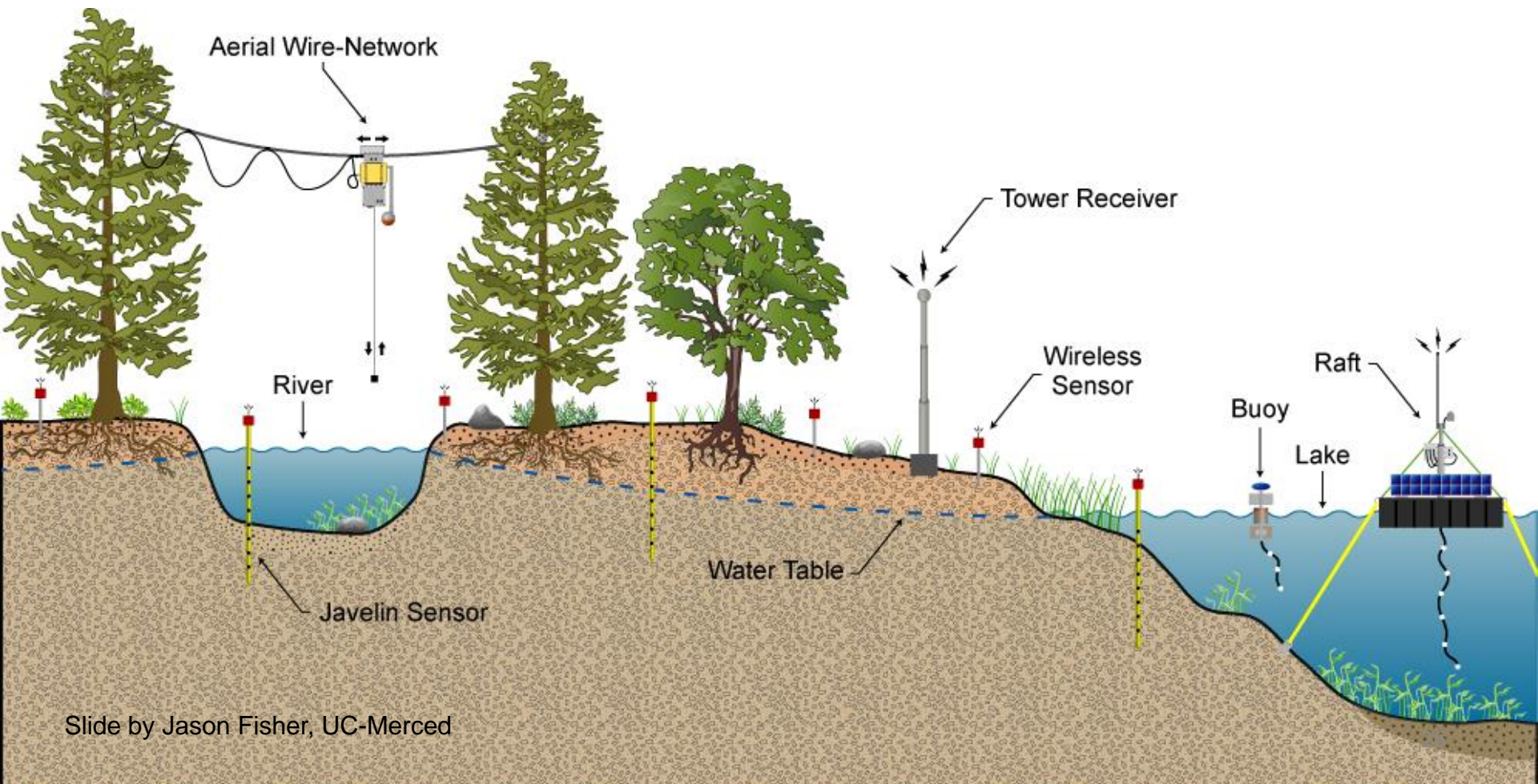
Click to rotate



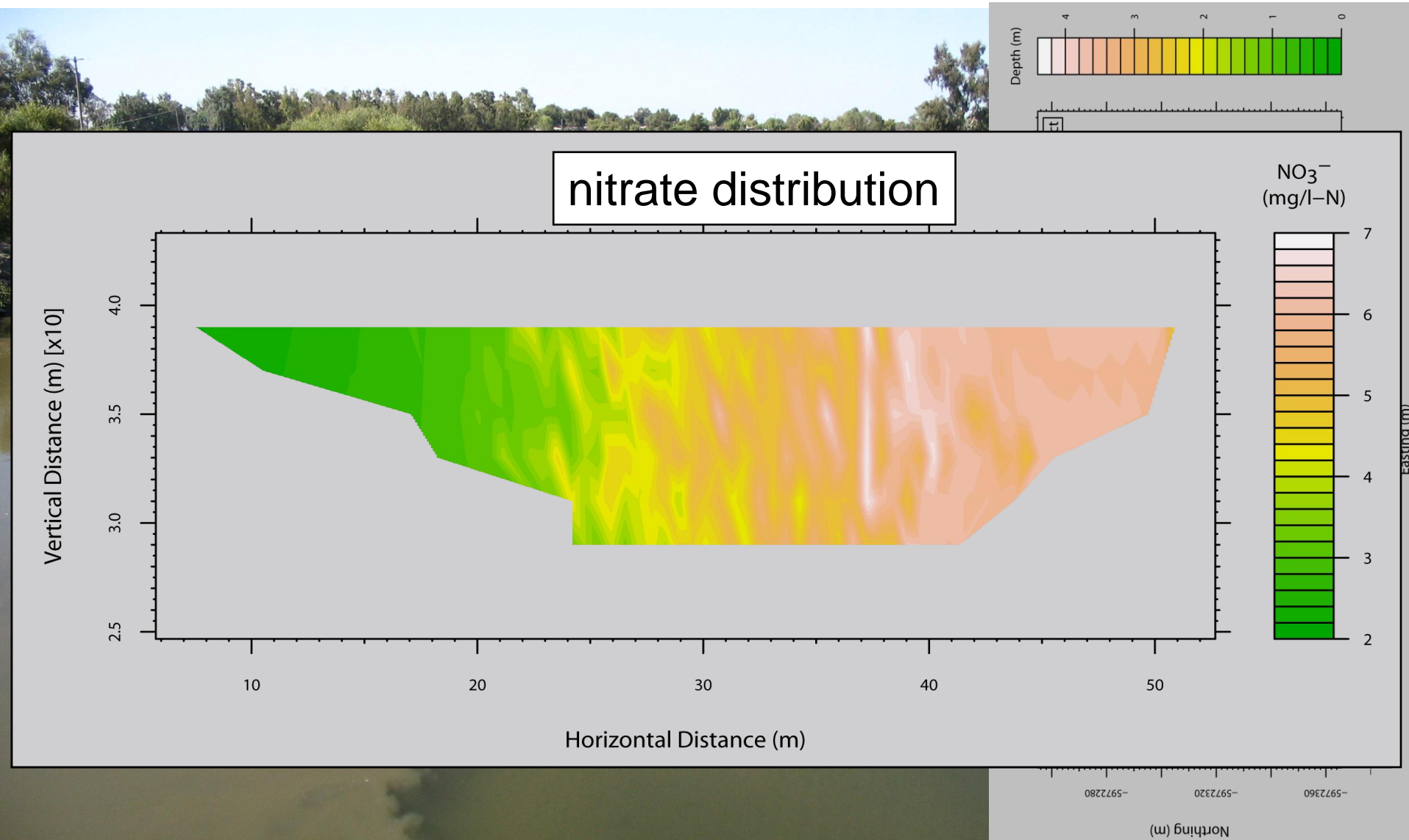
**Figure 3** | Schematic illustration of the dendrogram process. Shown is the

<sup>1</sup>Initiative in Innovative Computing at Harvard, Cambridge, Massachusetts 02138, USA. <sup>2</sup>Harvard-Smithsonian Center for Astrophysics, Cambridge, Massachusetts 02138, USA. <sup>3</sup>Department of Physics, University of British Columbia, Vancouver, Kelowna, British Columbia V1V 1V7, Canada. <sup>4</sup>Surgical Planning Laboratory and Department of Radiology, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts 02115, USA. <sup>5</sup>Present address: School of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts 02138, USA.

# Sensor networked science



# Sensor network data

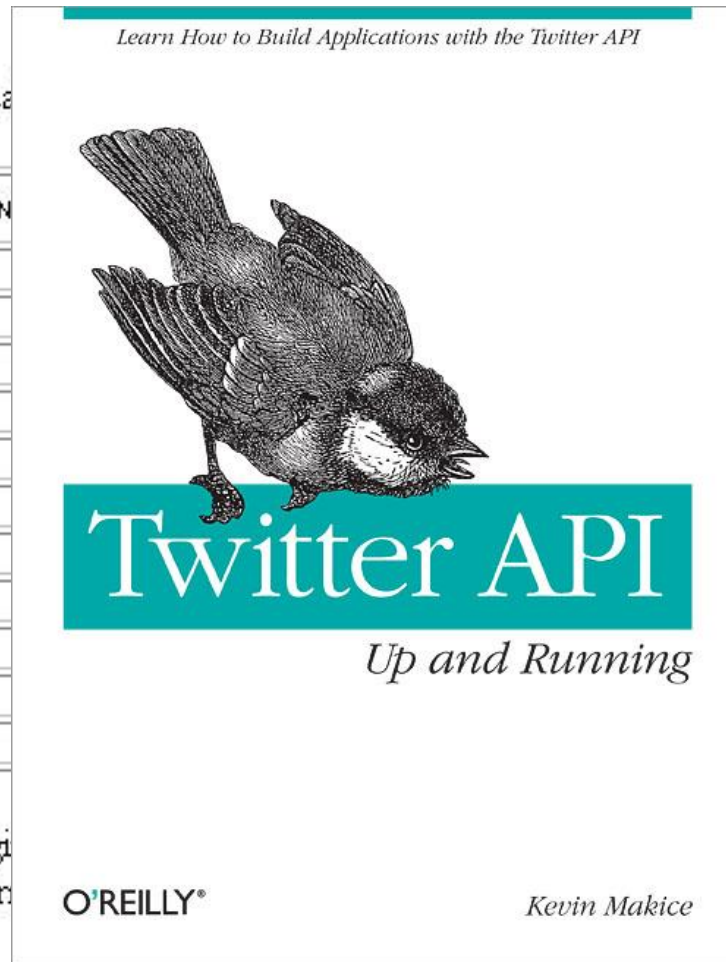


# Social science data

56. Generally speaking, do you usually vote for a Democrat, Republican, or Independent, or what?

RESPONSE	PUN
Strong Democrat	
Not very strong Democrat	
Independent, close to Democrat	
Independent (Neither, No response)	
Independent, close to Republican	
Not very strong Republican	
Strong Republican	
Other party, refused to say	
Don't know	
No answer	

See Appendix D: Recodes, for original response categories across surveys. If planning to perform analyses, see No. 56.



independent, or what?

	1994	1996	1998	COL: 240
7	423	400	370	6,046
1	644	577	597	8,756
0	341	356	349	4,581
5	369	457	477	4,882
8	282	258	244	3,379
9	519	500	484	6,265
0	321	307	239	3,479
7	44	43	63	530
0	0	0	0	10
9	49	6	9	188

Appendix N for changes in response categories across surveys. See GSS Methodological Report



CBETA 中

▼ CBETA 首頁

▼ 電子佛典集成光碟

▼ 下載

CBETA 首頁

熱門連結

Zacchetti, S. (2005). In Praise of the Light: A Critical Synoptic Edition with an Annotated Translation of Chapters 1-3 of Dharmarakṣa's Guang zan jing, Being the Earliest Chinese Translation of the Larger Prajnaparamita. Tokyo, Japan: The International Research Institute for Advanced Buddhology, Soka University. Retrieved from [http://iriab.soka.ac.jp/orc/Publications/BPPB/index\\_BPPB.html](http://iriab.soka.ac.jp/orc/Publications/BPPB/index_BPPB.html)



我們也有 專頁

追蹤

karmāntājivā  
virahitakuśalakāyavānmanas-  
karmāntājivāś<sup>109</sup> ca bhavanti sma

§ 1.66  
(147c 2b-148a  
0)

一切眾生得平等心，展轉相瞻如父、如母、如兄、如弟、如姊、如妹，各各同心，等無偏邪，皆行慈心。

PG 4r 4-5 (Ś 18, 22-19, 1; PD 10, 1-2; PSL *kā* a 4-5): sarvasatvāś ca sarvasatveṣu samacittā abhūvan\* yad uta<sup>110</sup> mātāpitr̥bhṛātr̥bhaginīsamacittāḥ mitrajñātisahāyasaamacittāḥ<sup>111</sup>

§ 1.67  
(148a 1-2)

一切群萌悉修十善，清淨梵行，無有塵埃。

PD 10, 2-3 (PG 4r 5-6; Ś 19, 2-3; PSL *kā* a 5): daśakuśalakarmapathasevinaś ca bhavanti sma<sup>112</sup> / brahmacāriṇaḥ śucayo nirāmayagandhāḥ<sup>113</sup>

§ 1.68  
(148a 2-4)

一切黎庶悉獲安隱，所得安隱猶如比丘得第三禪。于時衆生而致智慧，而悉具足善快調定，離於卑劣，速得和雅。

PG 4r 6-8 (PD 10, 3-8; PSL *kā* a 5-6; Ś 19, 3-8): sarvasatvās tasmin samaye sarvasukhasamarpitā abhūvan\* evaṃrūpeṇa sukhena samanvāgatāś<sup>114</sup> tadyathā (s) ṛṭṭiyadhyanasamāpannasya bhikṣoḥ sukhaṃ sarvasatvāś ca tasmin samaye evaṃrūpayā prajñayā samanvāgatā abhūvan\* yad evaṃ jānaṃti sma<sup>115</sup> • sādhu dānaṃ sādhu damaḥ sādhu saṃyamah<sup>116</sup> sādhu satyaṃ • sādhu apramādaḥ sādhu maitri sādhu karuṇā sādhu avihimsā prāṇibhūteṣu<sup>117</sup> •

<sup>110</sup> sarvasatvāś ... yad uta: not in PD & PSL.

<sup>111</sup> PG wrongly repeats verbatim this latter compound. PD 10, 2 and PSL have at this point a longer reading: mitramātyajñātisālohitasaamacittā. Note that Ś has all the words construed as one compound.

<sup>112</sup> PG 4r 5-6 & Ś 19, 2: daśakuśalakarma(tha)samanvāgatā [Ś without daśa-] abhūvan.

<sup>113</sup> PG 4r 6, Ś 19, 3 and PSL *kā* a 5: nirāmagandhāḥ, which seems to be the correct reading; after this word, PG & Ś + sarvakuśalavitarakavigatāḥ.

<sup>114</sup> PD 10, 4 & PSL *kā* a 5: idṛśaṃ sukhaṃ pratilabhante sma.

<sup>115</sup> yad ... sma: PD 10, 6 & PSL *kā* a 6: yad anyabuddhākṣetrasthā buddhā bhagavanta evaṃ [PSL + udānam] udānanti sma.

<sup>116</sup> Ś 19, 7: saṃyamah.

<sup>117</sup> sādhu dānaṃ ... prāṇibhūteṣu: PD 10, 7-8 & PSL *kā* a 6: sādhu damaḥ [PSL + sādhu śamaḥ] sādhu saṃyamah sādhu cirṇo brahmacāryāvāsaḥ sādhu prāṇibhūteṣu avihimseti.

登入 | Register

大藏經搜尋

## 論版主題

那落格 編輯  
10  
· 頭智巖的標

壯歲謁阿  
別傳之

那落格 編輯  
10  
!

閱讀全文

## 最新回應

1) 卷29錯字

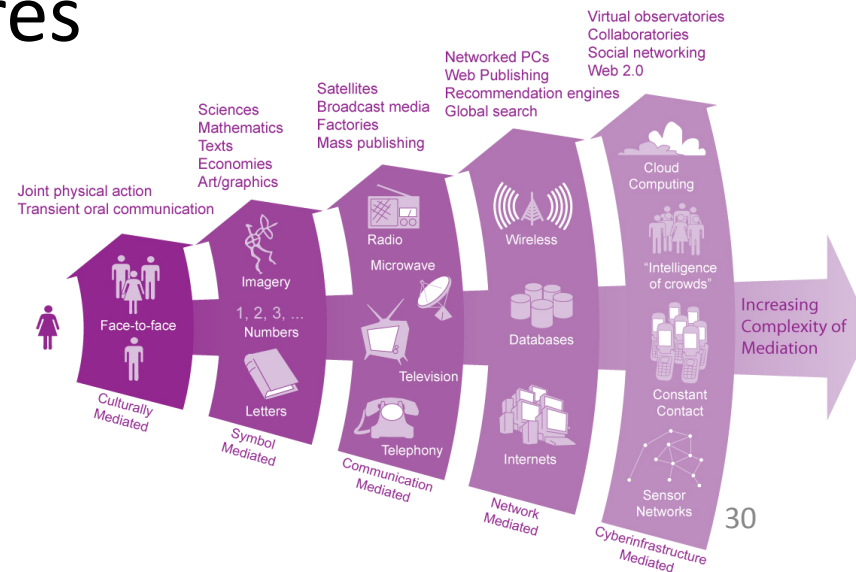
1) 牛頭智巖

· 就是梅  
· 也是花。不

29  
· 勝，呵呵。  
· 句也可改為

# Libraries, librarians, and data

1. Data are not publications
2. Data are not natural objects
3. Data are representations
4. Data sharing and reuse depend on knowledge infrastructures



# Data

- **Data are representations of observations, objects, or other entities used as evidence of phenomena for the purposes of research or scholarship.**



# Representations of research objects

**ORA**  
Oxford University Research Archive  
Research publications from the University of Oxford

SEARCH CONtribute BROWSE

Metastate Only SEARCH Help

**Top Downloads**

- Maxim David Bitting, (2012), Correction of misperceptions of justice in New Zealand and England. *CPHIL*, University of Oxford.
- John Barnes, The Sobolevsky VII foundation studies: an additional fragment. *Acta Orientalia*, 37, 15-20.
- John Barnes, The foundation studies of Sobolevsky VII. *Acta Orientalia*, 36, 38-54.
- Thomas Curtright, (2011), Coarse-grained modeling of DNA and DNA self-assembly. *CPHIL*, University of Oxford.
- Elisa Ripa, (2010), The reduction of children of immigrants in France. *CPHIL*, University of Oxford.

**Latest Additions**

- Nicole Gibry and Andrew Honey, (2012), The conservation of two corporate French manuscripts from the 16th century: two contrasting approaches?
- Debjyoti Bhattacharya, Shikharaj and Margaret Stevens, (2013), Are university admissions economically fair?
- Mark Irish, (2013), The synthesis and characterization of metal complexes containing chemically reduced bipyridyl ligand systems. *CPHIL*, University of Oxford.
- Andreas Pika, (2012), Essays in corporate restructuring, reputation and law. *CPHIL*, University of Oxford.
- Keith Hyman and Tina Fawcett, (2013-March/Apr), The ethics of carbon offsetting. *WIREs Climate Change*, 4(2), 11-66.

Follow @oxfordarch

The **Dataverse Network**™ Project

A Web Application for Sharing, Citing, Analyzing and Preserving Research Data

ABOUT SOFTWARE DATA MANAGEMENT GUIDES

A repository for research data that takes care of **long term preservation** and **good archival practices**, while researchers can **share, keep control of** and **get recognition** for their data.

Supports the sharing of open data and enables reproducible research.

**WHY**  
the  
Dataverse Network?

03:48 vimeo

figshare

Browse Upload

shareable

- Publish your negative data
- Quick and simple uploads
- All formats of research accepted

discoverable

- Secure and accessible
- Easy to manage your research data
- 1GB private space
- Unlimited public space

Find out more

Sign up

get credit for all your research

or store it privately for FREE\*

Find out more

\* up to 1GB of free space

Sign up

"Data which are determined to be a matter of public record and can therefore be made freely available, without restriction, are categorised as Public Access"

"Data must be properly curated throughout its life-cycle and released with the appropriate high-quality metadata."

"NERC is committed to supporting long-term environmental data management to enable continuing access to these data."

University of Hertfordshire

MRC Medical Research Council

NATURAL ENVIRONMENT RESEARCH COUNCIL

SEARCH Harvard Dataverse Network

TWEETS

THE DATA BLOG

SEARCH

The **Harvard Dataverse Network** is open to all researchers worldwide to share, cite, reuse and

**INEconomics** Six Ways to Separate Lies From #Statistics Lco(dobdTKWzq] via @BloombergView 22 hours 7 min ago.

Upcoming changes for version 3.5 of Dataverse Network (Subnetworks) May 2, 2013

SEARCH OUR SITE GO

**FIRST TIME HERE? HELPFUL INFORMATION**

**A QUICK GUIDE TO THE ARCHIVE**

2 of 10: We hold thousands of data collections for social science research and teaching, quantitative and qualitative

**WHO GIVES US DATA?**

Find out what kind of data are available to you

**LATEST NEWS & EVENTS**

Staff vacancy: Data and Services The UK Data Archive is seeking to appoint a Data ...

Get to know the UK Data Service Those who rely on social and economic data now ...

Service featured in LSE blog Two leaders from the UK Data Service appear as ...

**UK DATA SERVICE NOW ONLINE**

After months of planning, the newly formed UK Data Service launched its website on 21 March 2013.

READ MORE

What's new Data lifecycle Find data Who are we?

**DEPOSITING YOUR DATA**

We can help you find data for research and teaching with our catalogue of over 5,000 data collections

DEPOSIT DATA

OUR CATALOGUE

**OUR DATA IN USE**

Labour Force Survey: How does the health service distribute resources equitably?

OUR DATA IN USE

**OUR SERVICES**

UK Data Service data promoting evidence-based research

VIEW SITE

**CITING DATA**

PREPARING YOUR DATA

32

# Description and Classification



Research

Gen



FB2013\_03, released May 7th, 2013

A Database of *Drosophila* Genes & Genomes



The Open Biological and Biomedical Ontologies

Ontologies

Resources

Participate

About

The OBO Foundry is a collaborative experiment involving developers of science-based ontologies who are establishing a set of principles for ontology development with the goal of creating a suite of orthogonal interoperable reference ontologies in the biomedical domain. The groups developing ontologies who have expressed an interest in this goal are listed below, followed by other relevant efforts in this domain.

In addition to a listing of OBO ontologies, this site also provides a statement of the OBO Foundry principles, discussion fora, technical infrastructure, and other services to facilitate ontology development. We welcome feedback and encourage participation.

Click any column header to sort the table by that column. The link to the term request trackers for the listed ontologies.

## OBO Foundry ontologies

Title	Domain	Prefix	File	Last changed
<a href="#">Biological process</a>	biological process	GO	<a href="#">go.obo</a>	
<a href="#">Cellular component</a>	anatomy	GO	<a href="#">go.obo</a>	
<a href="#">Chemical entities of biological interest</a>	biochemistry	CHEBI	<a href="#">chebi.obo</a>	
<a href="#">Molecular function</a>	biological function	GO	<a href="#">go.obo</a>	
<a href="#">Phenotypic quality</a>	phenotype	PATO	<a href="#">quality.obo</a>	
<a href="#">PRotein Ontology (PRO)</a>	proteins	PR	<a href="#">pro.obo</a>	
<a href="#">Xenopus anatomy and development</a>	anatomy	XAO	<a href="#">xenopus_anatomy_edit.obo</a>	
<a href="#">Zebrafish anatomy and development</a>	anatomy	ZFA	<a href="#">zebrafish_anatomy.obo</a>	2013/04/12

## OBO Foundry candidate ontologies and other ontologies of interest

Title	Domain	Prefix	File	Last changed
<a href="#">Adverse Event Reporting Ontology</a>	health	AERO	<a href="#">aero.owl</a>	
<a href="#">Amphibian gross anatomy</a>	anatomy	AAO	<a href="#">AAO_v2_edit.obo</a>	
<a href="#">Amphibian taxonomy</a>	anatomy	ATO	<a href="#">amphibian_taxonomy.obo</a>	
<a href="#">Anatomical Entity Ontology</a>	anatomy	AEO	<a href="#">aao.obo</a>	2012/06/01
<a href="#">Ascomycete phenotype ontology</a>	phenotype	APO	<a href="#">ascomycete_phenotype.obo</a>	2013/05/02
<a href="#">Basic Formal Ontology</a>	upper	BFO	<a href="#">1.1</a>	
<a href="#">Bilateria anatomy</a>	anatomy	BILA	<a href="#">bilateria_mrca.obo</a>	
<a href="#">Biological imaging methods</a>	experiments	FBbi	<a href="#">image.obo</a>	2011/05/24
<a href="#">BRENDA tissue / enzyme source</a>	anatomy	BTO	<a href="#">BrendaTissueOBO</a>	
<a href="#">C. elegans development</a>	anatomy	WBle	<a href="#">worm_development.obo</a>	

### Quick Links

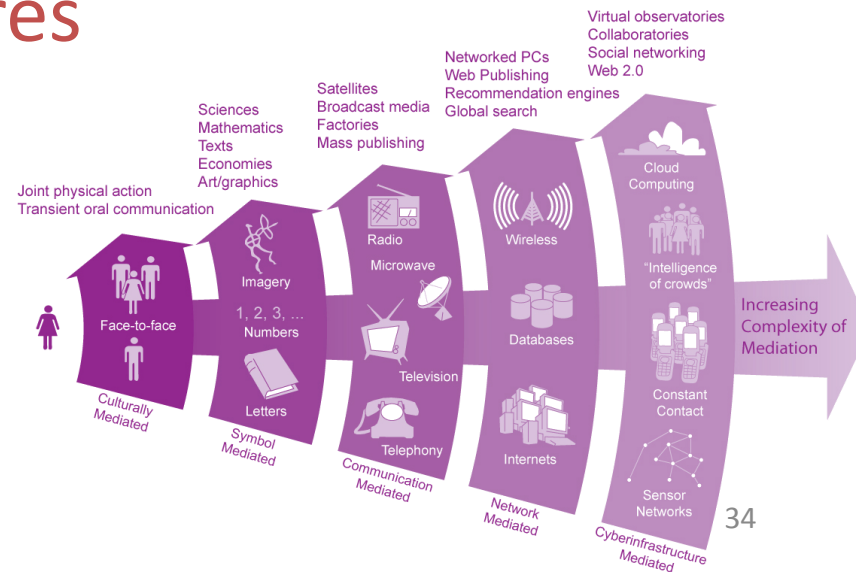
- [Mappings between ontologies](#)
- [Download alternate formats](#)
- [About the OBO Foundry](#)
- [Current events](#)
- [How to join](#)
- [OBO Foundry paper in Nature Biotechnology](#)  
November 2007

### Other Ontology Lists

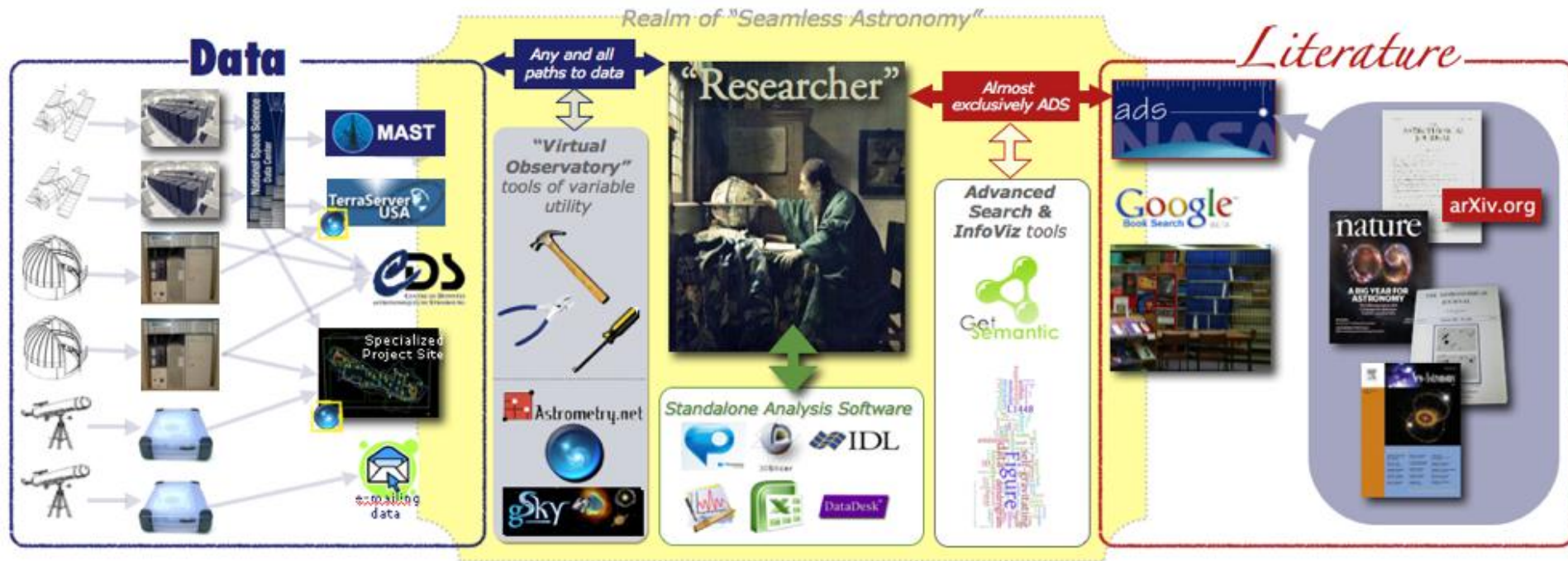
- [BioPortal](#) (NCBO's ontology repository)
- [Ontology Lookup Service \(OLS\)](#) (term lookup)

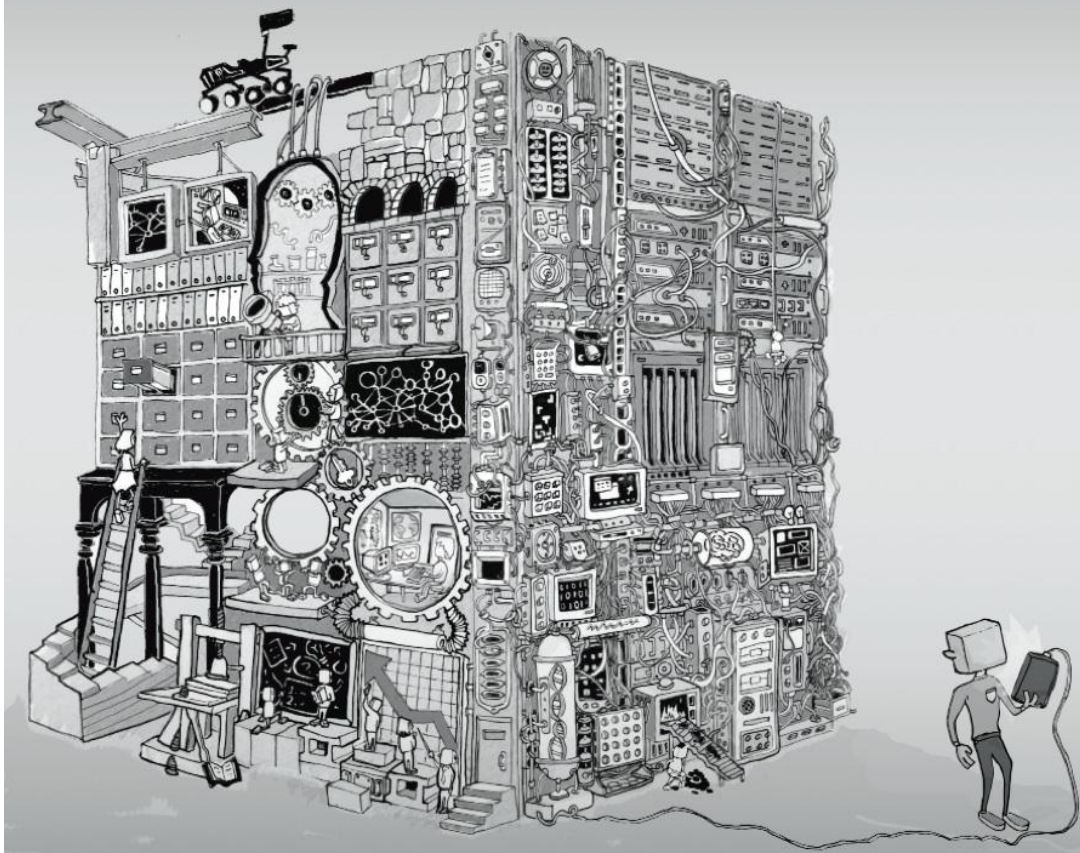
# Libraries, librarians, and data

1. Data are not publications
2. Data are not natural objects
3. Data are representations
4. Data sharing and reuse depend on knowledge infrastructures



# Knowledge Infrastructures

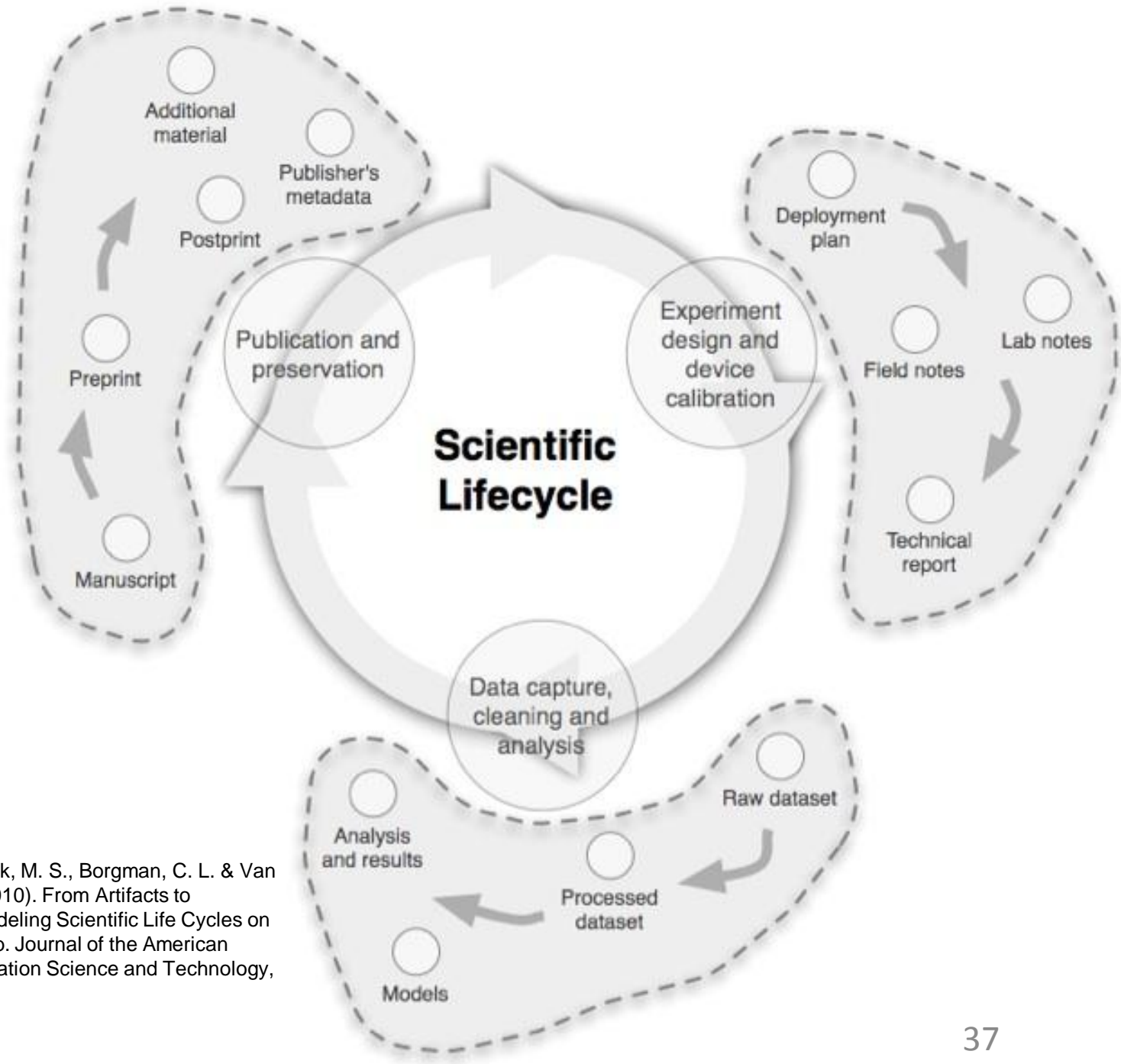




Knowledge Infrastructures:  
Intellectual Frameworks and Research Challenges

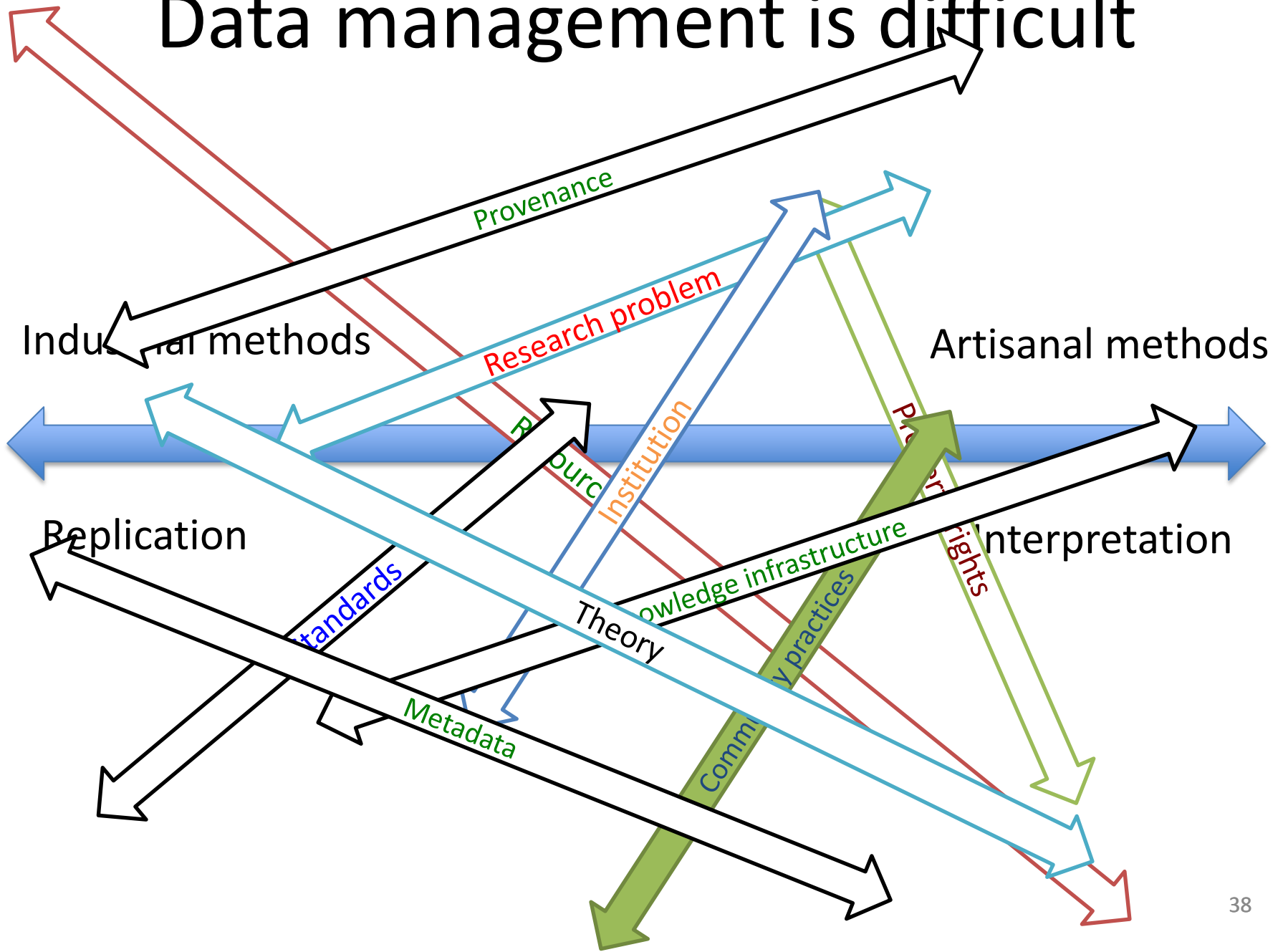
*Report of a workshop sponsored by the National Science Foundation and the Sloan Foundation  
University of Michigan School of Information, 25-28 May 2012*

<http://knowledgeinfrastructures.org>



Pepe, A., Mayernik, M. S., Borgman, C. L. & Van de Sompel, H. (2010). From Artifacts to Aggregations: Modeling Scientific Life Cycles on the Semantic Web. *Journal of the American Society for Information Science and Technology*, 61(3): 567–582.

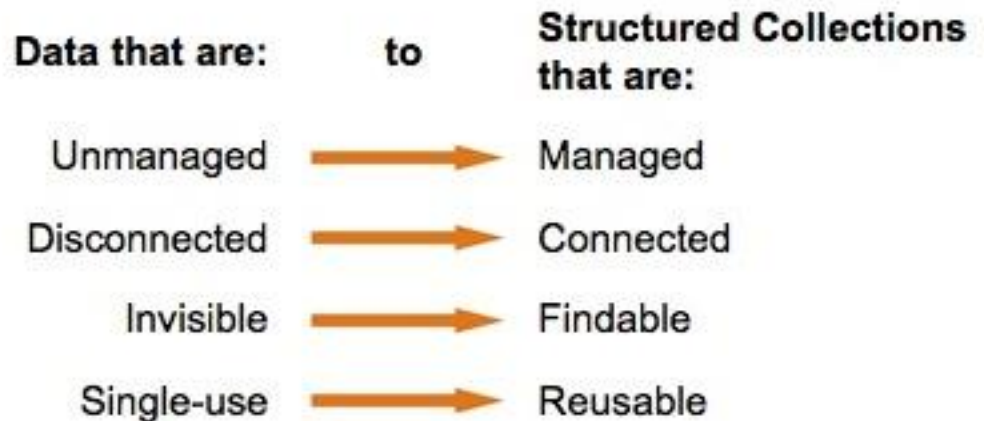
# Data management is difficult



# Ways to Release Data

- Contribute to archive
- Attach to journal article
- Post on local website
- License on request
- Release on request
- ....

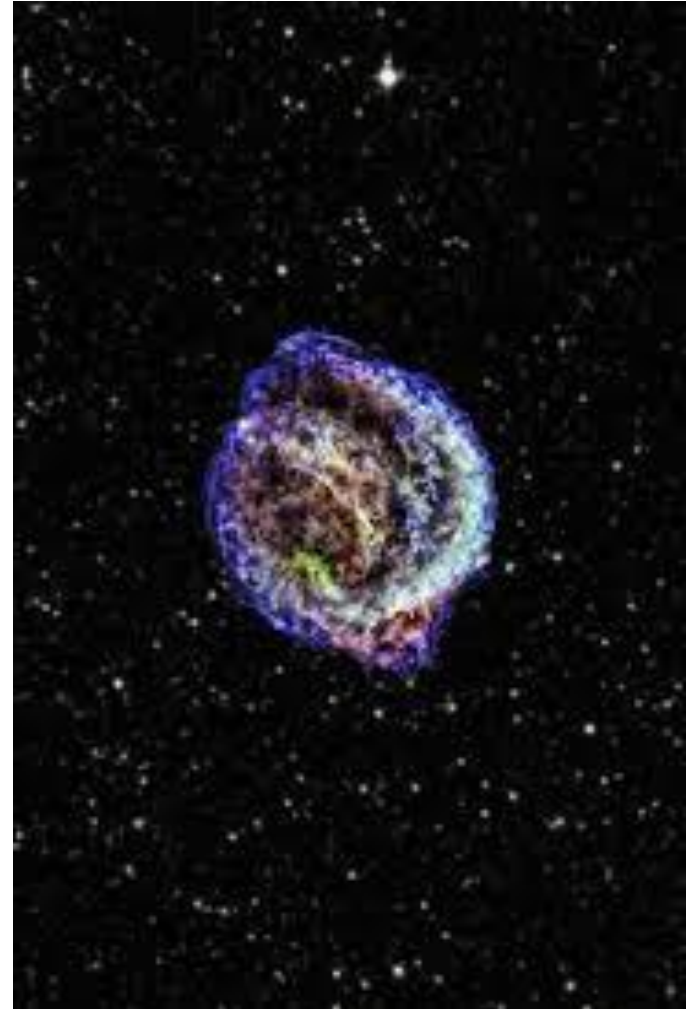
## ANDS goals





# Degrees of reuse

- Reuse by investigator
- Reuse by collaborators
- Reuse by colleagues
- Reuse by unaffiliated others
- Reuse at later times
  - Months
  - Years
  - Decades
  - Centuries



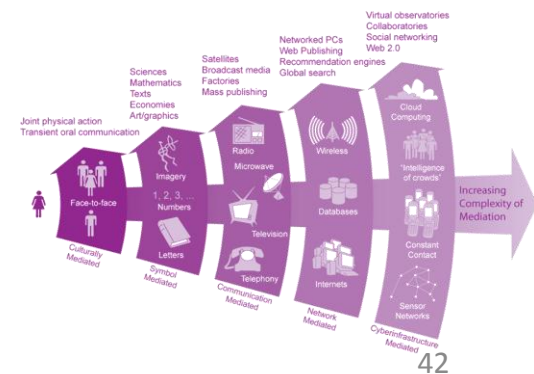
# Economics of the Knowledge Commons

		Subtractability / Rivalry	
		Low	High
Exclusion	Difficult	<b>Public Goods</b> General knowledge Public domain data	<b>Common-pool resources</b> Libraries Data archives
	Easy	<b>Toll or Club Goods</b> Subscription journals Subscription data	<b>Private Goods</b> Printed books Raw or competitive data

Adapted from C. Hess & E. Ostrom (Eds.), *Understanding knowledge as a commons: from theory to practice*. Cambridge, Mass.: MIT Press.

# Conclusions

- Stakeholders
- Economics of research data
- Sustainability
- Knowledge infrastructures and libraries



# Stakeholders in research data

- Scholars

- Grants and funding
- Credit and promotion
- Publishing
- Managing data

- Others

- Scholars
- Students
- Readers / public at large
- Universities
- Funding agencies
- Repositories
- Governments
- Private enterprise
- Libraries
- Archives
- Museums...



# Economics of Research Data

- Private goods
- Common-pool resources
- Toll or club goods
- Public goods



# Sustainability

- What to keep
- Why to keep them
- How to keep them
- How long to keep them
- Who will govern them
- What kinds of expertise are required



# Knowledge Infrastructure and Libraries

- Technical fabric
  - Data archives
  - Metadata, provenance, classification, collections
  - Tools, services, support
- Social fabric
  - Data curation workforce
  - Data management training
- Trust fabric
  - Memory institutions
  - Stewardship



# Acknowledgements

---

- UCLA Data Practices team
  - Rebekah Cummings, Peter Darch, David Fearon, Ariel Hernandez, Matthew Mayernik, Alberto Pepe, Ashley Sands, Katie Shilton, Sharon Traweek, Jillian Wallis, Laura Wynholds, Kan Zhang
- Research funding
  - National Science Foundation
  - Alfred P. Sloan Foundation
  - Microsoft External Research



**Microsoft®**

- University of Oxford
  - Balliol College
  - Oliver Smithies Fellowship
  - Oxford Internet Institute
  - Oxford eResearch Center
  - Bodleian Library





# Provocations

1. How do rights, responsibilities, and risks associated with research data vary by discipline and by stakeholder?
2. How can data be exchanged across domains, contexts, and over time?
3. How do publications and data differ?
4. What are scholars' motivations to share their data?
5. What kinds of expertise are required to manage, curate, and sustain access to research data?
6. How can knowledge infrastructures adapt to the needs of generations of scholars and the demands of competing stakeholders?

# 10 Simple Rules for the Care and Feeding of Scientific Data\*

1. Love your data, and help others love it too.
2. Share your data online, with a permanent identifier.
3. Conduct science with a particular level of reuse in mind.
4. Publish workflow as context
5. Link your data to your publications as early as possible.
6. Publish your code (even the small bits).
7. Say how you want to get credit for your data (and software).
8. Foster and use data repositories.
9. Reward colleagues who share their data properly.
10. Be a booster for data science.