Title
Data Scholarship in the Humanities

Permalink
https://escholarship.org/uc/item/6xv4b7bd

Author
Borgman, Christine L.

Publication Date
2014-10-09
Data Scholarship in the Humanities

Christine L. Borgman

Professor and Presidential Chair in Information Studies
University of California, Los Angeles
Visiting Professor KNAW, DANS – Digital Archiving and Data Services

New Trends in eHumanities, Meertens Institute, Amsterdam (eHumanities Group Research Meeting)
Royal Netherlands Academy of Arts and Sciences 9 October 2014

Gustave Dore, Rime of the Ancient Mariner, Woodcut, 1798
Day after day, day after day,
We stuck, nor breath nor motion;
As idle as a painted ship
Upon a painted ocean.

Water, water, every where,
And all the boards did shrink;
Water, water, every where,
Nor any drop to drink.

Stanzas from
_The Rime of the Ancient Mariner_
Samuel Taylor Coleridge, 1798
Big Data, Little Data, No Data: Scholarship in the Networked World*

- Part I: Data and Scholarship
  - Ch 1: Provocations
  - Ch 2: What Are Data?
  - Ch 3: Data Scholarship
  - Ch 4: Data Diversity
- Part 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
- Part 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

*C. L. Borgman (2015, January) MIT Press
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.

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
  - 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
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 unreliable investigations
- Reduction of duplicated effort
MEMORANDUM FOR THE HEADS OF EXECUTIVE DEPARTMENTS AND AGENCIES

FROM: John P. Holdren
       Director

SUBJECT: Increasing Access to the Results of Federally Funded Scientific Research

1. Policy Principles

The Administration is committed to ensuring that, to the greatest extent and with the fewest constraints possible and consistent with law and the objectives set out below, the direct results of federally funded scientific research are made available to and useful for the public, industry, and the scientific community. Such results include peer-reviewed publications and digital data.

Scientific research supported by the Federal Government catalyzes innovative breakthroughs that drive our economy. The results of that research become the grist for new insights and are assets for progress in areas such as health, energy, the environment, agriculture, and national security.
Chinese science gets mass transformation

Teamwork at centre of Chinese Academy of Sciences reform.

David Cyranoski

23 September 2014

BEIJING
Precondition:

Researchers share data
Data sharing

• Scholars’ perspectives
  – Rewards
  – Responsibility
  – Data
  – Incentives

• Open data

• Knowledge infrastructure

Persistent URL: photography.si.edu/SearchImage.aspx?id=5799
Repository: Smithsonian Institution Archives
Data sharing

• Scholars’ perspectives
  – Rewards
  – Responsibility
  – Data
  – Incentives
• Open data
• Knowledge infrastructure

Persistent URL: photography.si.edu/SearchImage.aspx?id=5799
Repository: Smithsonian Institution Archives
Rewards

- Publications
- Publications
- Publications
- Publications
- Publications
- Grants
- Awards and honors
- Teaching
- Service
- Data

http://blog.startfreshtoday.com/Portals/170402/images/improve-credit-score1.jpg
Functions of Scholarly Publications

• Legitimization
  – Authority, quality
  – Priority, trustworthiness

• Dissemination
  – Awareness
  – Diffusion
  – Publicity

• Access, preservation, curation
  – Availability
  – Discovery
  – Retrieval
  – Persistence

Data sharing

• Scholars’ perspectives
  – Rewards
  – Responsibility
  – Data
  – Incentives

• Open data

• Knowledge infrastructure

Persistent URL: photography.si.edu/SearchImage.aspx?id=5799
Repository: Smithsonian Institution Archives
Responsibility

Publications are arguments made by authors, and data are the evidence used to support the arguments.

Responsibility

- Publications
  - Independent units
  - Authorship is negotiated

- Data
  - Compound objects
  - Ownership is rarely clear
  - Attribution
    - Long term responsibility: Investigators
    - Expertise for interpretation: Data collectors and analysts
Attribution of data

• Legal responsibility
  – Licensed data
  – Specific attribution required

• Scholarly credit: contributorship
  – “Author” of data
  – Contributor of data to this publication
  – Colleague who shared data
  – Software developer
  – Data collector
  – Instrument builder
  – Data curator
  – Data manager
  – Data scientist
  – Field site staff
  – Data calibration
  – Data analysis, visualization
  – Funding source
  – Data repository
  – Lab director
  – Principal investigator
  – University research office
  – Research subjects
  – Research workers, e.g., citizen science...

Data sharing

• Scholars’ perspectives
  – Rewards
  – Responsibility
  – Data
  – Incentives

• Open data

• Knowledge infrastructure

Persistent URL: photography.si.edu/SearchImage.aspx?id=5799
Repository: Smithsonian Institution Archives
What are data?

Pisa Griffin
Arte islamica, ippogrifo, XI sec 03, own work

Marie Curie's notebook aip.org

hudsonalpha.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

ncl.ucar.edu

NASA Astronomy Picture of the Day
http://onlineqda.hud.ac.uk/Intro_QDA/Examples_of_Qualitative_Data.php

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. E5 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 Süleymani. Goes in and out once a day. He gets 8,000 a month. Zafer then said, keskin dolmuş (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.
A role for self-gravity at multiple length scales in the process of star formation

Alyssa A. Goodman, Erik W. Rosolowsky, Michelle A. Borkin, Jonathan B. Foster, Michael Halle, Jens Kauffmann & Jaime E. Pineda

Self-gravity plays a decisive role in the final stages of star formation, where dense cores (size ~ 0.1 parsecs) inside molecular clouds collapse to form star plus disk systems. But self-gravity's role at earlier times (and on larger length scales, such as ~ 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. 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 12CO 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 natural 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 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 of parsecs across, and surrounded by atomic gas) to be marginally self-gravitating. 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. 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 CLUMPJND. In three-dimensional (3D) spectral-line data cubes, CLUMPJND 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 CLUMPJND decomposition of it based on 12CO 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 attributed 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 contours of the color image correspond to the near-infrared bands J (blue), H (green) and K (red), and the contours of integrated intensity are from 12CO (1–0) emission. Integrated intensity is monochromatic, but not quite linearly (see Supplementary Information), related to column density, 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 protostars 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 'Clouds'. The four yellow box 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. Astronomers show the location 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.
Center for Embedded Networked Sensing

- NSF Science & Tech Ctr, 2002-2012
- 5 universities, plus partners
- 300 members
- Computer science and engineering
- Science application areas

Slide by Jason Fisher, UC-Merced, Center for Embedded Networked Sensing (CENS)
Borgman, et al. (2007). Drowning in data: Digital library architecture to support scientific use of embedded sensor networks. JCDL
Documenting Data for Interpretation

Engineering researcher: “Temperature is temperature.”

Biologist: “There are hundreds of ways to measure temperature. ‘The temperature is 98’ is low-value compared to, ‘the temperature of the surface, measured by the infrared thermopile, model number XYZ, is 98.’ That means it is measuring a proxy for a temperature, rather than being in contact with a probe, and it is measuring from a distance. The accuracy is plus or minus .05 of a degree. I [also] want to know that it was taken outside versus inside a controlled environment, how long it had been in place, and the last time it was calibrated, which might tell me whether it has drifted..”

CENS Robotics team
Center for Dark Energy Biosphere Investigations

International Ocean Discovery Program
lodp.tamu.org

- NSF Science & Tech Ctr, 2010-2020
- 20 universities, plus partners (35 institutions)
- 90 scientists
- Biological sciences
- Physical sciences

Repository for seafloor cores. Photo: Peter Darch
Social science data

56. Generally speaking, do you usually think of yourself as a Republican, Democrat, Independent, or what?

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Strong Democrat</td>
<td>0</td>
<td>2197</td>
<td>143</td>
<td>1271</td>
<td>151</td>
<td>864</td>
<td>227</td>
<td>423</td>
<td>400</td>
<td>370</td>
<td>6,046</td>
</tr>
<tr>
<td>Not very strong Democrat</td>
<td>1</td>
<td>3482</td>
<td>109</td>
<td>1655</td>
<td>89</td>
<td>1,282</td>
<td>321</td>
<td>644</td>
<td>577</td>
<td>597</td>
<td>8,756</td>
</tr>
<tr>
<td>Independent, close to Democrat</td>
<td>2</td>
<td>1768</td>
<td>44</td>
<td>904</td>
<td>51</td>
<td>578</td>
<td>190</td>
<td>341</td>
<td>356</td>
<td>349</td>
<td>4,581</td>
</tr>
<tr>
<td>Independent (Neither, No response)</td>
<td>3</td>
<td>1736</td>
<td>30</td>
<td>855</td>
<td>32</td>
<td>721</td>
<td>205</td>
<td>369</td>
<td>457</td>
<td>477</td>
<td>4,882</td>
</tr>
<tr>
<td>Independent, close to Republican</td>
<td>4</td>
<td>1106</td>
<td>8</td>
<td>743</td>
<td>9</td>
<td>571</td>
<td>158</td>
<td>282</td>
<td>258</td>
<td>244</td>
<td>3,379</td>
</tr>
<tr>
<td>Not very strong Republican</td>
<td>5</td>
<td>2011</td>
<td>8</td>
<td>1259</td>
<td>15</td>
<td>1,170</td>
<td>299</td>
<td>519</td>
<td>500</td>
<td>484</td>
<td>6,265</td>
</tr>
<tr>
<td>Strong Republican</td>
<td>6</td>
<td>1009</td>
<td>8</td>
<td>751</td>
<td>2</td>
<td>662</td>
<td>180</td>
<td>321</td>
<td>307</td>
<td>239</td>
<td>3,479</td>
</tr>
<tr>
<td>Other party, refused to say</td>
<td>7</td>
<td>243</td>
<td>0</td>
<td>75</td>
<td>1</td>
<td>44</td>
<td>17</td>
<td>44</td>
<td>43</td>
<td>63</td>
<td>530</td>
</tr>
<tr>
<td>Don’t know</td>
<td>8</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>No answer</td>
<td>9</td>
<td>64</td>
<td>4</td>
<td>29</td>
<td>3</td>
<td>15</td>
<td>9</td>
<td>49</td>
<td>6</td>
<td>9</td>
<td>188</td>
</tr>
</tbody>
</table>

See Appendix D: Recodes, for original question format and method of recoding. See Appendix N for changes across surveys. If planning to perform trend analysis with this variable, please consult GSS Methodological Report No. 56.

http://dss.princeton.edu/images/gss.gif
Social science data

See Appendix D: Recodes, for original coding differences across surveys. If planning to perform calculations across the years, see No. 56.

http://dss.princeton.edu/images/gss.gif
CLAROS Data

The CLAROS Data service provides a RESTful interface for the data of the CLAROS Project, and complements the CLAROS Explorer.

This service provides metadata about archaeology and art in machine-readable formats such as RDF, JSON and KML. The data for the CLAROS project are modelled using the Erlangen OWL-DL 1.0 implementation of the CIDOC Conceptual Reference Model using the http://purl.org/NET/crm-owl# namespace.

You can use the Objects and People views to start exploring, or try your hand at a SPARQL query.

Hosted by the University of Oxford's e-research centre, OeRC
This interface was built using Fuseki, humfrey, and a tweaked frontend, all open-source software.

http://data.clarosnet.org/
The aim of this project is to perform a comparative study of three artworks (bronze casts of Islamic provenance), to discover evidence of similarities and to get new insight on their origin.

Probably produced within the Islamic Mediterranean in the eleventh century, the Griffin has incised on its body a long inscription in Arabic expressing good wishes. Captured by the Pisans, it underwent an extraordinary transformation: for centuries it was a terrifying, sound-producing guardian figure on top of the roof of Pisa Cathedral. The present project is focused on the Griffin but also includes alongside it other bronze animal sculptures such as a Lion and a Falcon. It is hoped that the interdisciplinary study of the Griffin will shed light on the significance of such objects in a global Mediterranean culture.

Videos

The Pisa Griffin: an introduction

http://vcg.isti.cnr.it/griffin/

Arte islamica, ippogrifo, XI sec 03, own work
Brick inscribed with the Sutra on Dependent Origination **Gorakhpur district, late 5th century - early 6th century AD. Ashmolean Museum**
Humanities data

Brick inscribed with the Sutra on Dependent Origination
Gorakhpur district, late 5th century - early 6th century AD.
Ashmolean Museum
Data are representations of observations, objects, or other entities used as evidence of phenomena for the purposes of research or scholarship.

Data sharing

• Scholars’ perspectives
  – Rewards
  – Responsibility
  – Data
  – Incentives

• Open data

• Knowledge infrastructure

Persistent URL: photography.si.edu/SearchImage.aspx?id=5799
Repository: Smithsonian Institution Archives
Incentives

• Publications that report the research
Vs.
• Data that are reusable by others

Image: Alyssa Goodman, Harvard Astronomy
Metadata

• Metadata is structured information that describes, explains, locates, or otherwise makes it easier to retrieve, use, or manage an information resource.*
  – descriptive
  – structural
  – administrative

*National Information Standards Organization 2004
Provenance

- Libraries: Origin or source
- Museums: Chain of custody
- Internet: Provenance is information about entities, activities, and people involved in producing a piece of data or thing, which can be used to form assessments about its quality, reliability or trustworthiness.*

*World Wide Web Consortium (W3C) Provenance working group

British Library, provenance record: Bestiary - caption: 'Owl mobbed by smaller birds'
Reuse across place and time

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

http://chandra.harvard.edu/photo/2013/kepler/kepler_525.jpg
Data sharing

• Scholars’ perspectives
  – Rewards
  – Responsibility
  – Data
  – Incentives

• Open data

• Knowledge infrastructure

Persistent URL: photography.si.edu/SearchImage.aspx?id=5799
Repository: Smithsonian Institution Archives
Open Data - 1

• 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

<table>
<thead>
<tr>
<th>Research Area</th>
<th>Status</th>
<th>Round 1</th>
<th>Round 2</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Application Scientists</strong></td>
<td>Faculty</td>
<td>7</td>
<td>6</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>Staff</td>
<td>5</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Student</td>
<td>3</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td><strong>Technology Researchers</strong></td>
<td>Faculty</td>
<td>4</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Staff</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Student</td>
<td>2</td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td><strong>Totals</strong></td>
<td></td>
<td>22</td>
<td>21</td>
<td>43</td>
</tr>
</tbody>
</table>

doi:10.1371/journal.pone.0067332.t001
Open Data - 2

- Data that meets the criteria of intelligent openness. Data must be accessible, usable, assessable and intelligible.


*Indian Girl Pitching up Haystack*
Uploaded by The U.S. National Archives on December 9, 2009
SCHUMACHER, Ernst Friedrich

Small is beautiful: Economics as if people mattered.
Image from a photo album (AL-27) which belonged to Ray Fife, a pioneer aviation mechanic who built a complete plane at age 16. In 1963 he built a reproduction Curtiss Pusher Biplane. This album documents the building of the pusher and also contains photos used to research the project.
Open Data - 3

• Openness, flexibility, transparency, legal conformity, protection of intellectual property, formal responsibility, professionalism, interoperability, quality, security, efficiency, accountability, and sustainability.

Discovery and Interpretation

- Identify the form and content
- Identify related objects
- Interpret
- Evaluate
- Open
- Read
- Compute upon
- Reuse
- Combine
- Describe
- Annotate...

Image from Soumitri Varadarajan blog. Iceberg image © Ralph A. Clevenger. Flickr photo
Data sharing

- Scholars’ perspectives
  - Rewards
  - Responsibility
  - Data
  - Incentives
- Open data
- Knowledge infrastructure

Persistent URL: [photography.si.edu/SearchImage.aspx?id=5799](photography.si.edu/SearchImage.aspx?id=5799)
Repository: Smithsonian Institution Archives
Day after day, day after day,  
We stuck, nor breath nor motion;  
As idle as a painted ship  
Upon a painted ocean.

Water, water, every where,  
And all the boards did shrink;  
Water, water, every where,  
Nor any drop to drink.

_Stanzas from_  
_The Rime of the Ancient Mariner_  
Samuel Taylor Coleridge, 1798
Emerging themes in data practices

- Scarcity or abundance of data
- Centrality of data to research
- Time frame of research
- Heterogeneity of expertise
- Maturity of standards
- Community building

# Economics of the Knowledge Commons

<table>
<thead>
<tr>
<th>Subtractability / Rivalry</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
</tr>
<tr>
<td>High</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Exclusion</th>
<th>Difficult</th>
<th>Public Goods</th>
<th>Common-pool resources</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>General knowledge</td>
<td>Libraries</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Public domain data</td>
<td>Data archives</td>
</tr>
<tr>
<td>Easy</td>
<td></td>
<td>Toll or Club Goods</td>
<td>Private Goods</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Subscription journals</td>
<td>Printed books</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Subscription data</td>
<td>Raw or competitive data</td>
</tr>
</tbody>
</table>

Adapted from C. Hess & E. Ostrom (Eds.), *Understanding knowledge as a commons: From theory to practice*. MIT Press.
To share data, scholars need

- Fresh water
  - Tools
  - Services
  - Skills
  - Resources
  - Incentives

http://environment.nationalgeographic.com/environment/habitats/water-pressure/
To share data, scholars need

- Life boats
  - Repositories
  - Governance models
  - Provenance models
  - Data stewardship workforce
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
Acknowledgements

UCLA Data Practices team

- Peter Darch, Milena Golshan, Irene Pasquetto, Ashley Sands, Sharon Traweek
- Former members: Rebekah Cummings, David Fearon, Ariel Hernandez, Elaine Levia, Jaklyn Nunga, Matthew Mayernik, Alberto Pepe, Kalpana Shankar, Katie Shilton, Jillian Wallis, Laura Wynholds, Kan Zhang

- Research funding: National Science Foundation, Alfred P. Sloan Foundation, Microsoft Research, DANS-Netherlands
- University of Oxford: Balliol College, Oliver Smithies Fellowship, Oxford Internet Institute, Oxford eResearch Center, Bodleian Library