

UCLA

Publications

Title

Syllabus for Data Curation and Policy, Part II, Spring 2015, UCLA Information Studies

Permalink

<https://escholarship.org/uc/item/4kz1w060>

Author

Borgman, Christine L.

Publication Date

2015

Data Curation and Policy

Winter, 2015, UCLA Information Studies 262B,
Weds, 9am-12:20pm, IS Room 245,
Christine L. Borgman, Professor & Presidential Chair
235 GSE&IS bldg, 310-825-6164; Christine.Borgman@ucla.edu

Course Description: 262A, 262B	1
Course Objectives	2
Course Materials	2
Office Hours	2
Grading	2
Summary of Assignment due dates	3
Topics, Readings, and Guest Speakers	3
Week 1, April 1: Selection and appraisal.....	3
Week 2, April 8: Data management in the field/data management in the university	4
Week 3, April 15: Technologies of persistence and identification	5
Week 4, April 22: Data, Publishing, and Publishers: 8am-11:20am session	6
Week 5, April 29: Economics of data: preservation, access, and sustainability	7
Week 6, May 6: Provenance in data: WWW and archival approaches, Paul Groth, Free University of Amsterdam and Elsevier Research, by video	7
Week 7, May 13: Ethics in use of public data for research Prof Katie Shilton, U of Maryland, and Jeffrey Burke, UCLA	8
Week 8, May 20: Intellectual property in data: Peter Hirtle, Berkman Center, Harvard, by video	8
Week 9: choice of two topics:.....	8
Week 9, May 25: University-Industry partnerships: Data transfer	8
Week 9, May 25: Data in public engagement	9
Week 10, June 3: Project presentations	9
Week 11 (Exam Week), June 8 (Monday): Projects due	9
References	10

Course Description: 262A, 262B

Data are both process and products of the research enterprise. Increasingly, data are viewed as scholarly products to be managed, shared, and reused. Funding agencies are requiring data management plans as part of grant proposals, journals are requiring the release of data to reviewers and readers alike, and libraries and archives are adding data to their collections. Managing data is a complex process, involving expertise in technology, knowledge organization, information policy, and in the research domain.

These two courses (winter and spring) survey the landscape of data management, practices, services, and policy, including the uses of data in the sciences, social sciences, and humanities; data management practices (e.g., metadata, provenance, technical standards); national and international data policy (e.g., intellectual property, release policies, open access, economics); management of data by research teams, data centers, libraries, and archives; and data curation, preservation, and stewardship. The courses are intended for graduate students in information studies and in any domain that requires the management of research data. By bringing together students from across campus, these seminar courses will engage students in practical, professional, and theoretical challenges in the use and reuse of research data. Assignments include analyses of data archives, data management plans, curating data for a research team, and domain-specific activities. Students will work in teams with UCLA researchers and will make class presentations. This is a two-term course. Students taking Part I (winter) are not required to take Part II (spring), but Part I is pre-requisite to Part II.

Course Objectives

1. Students will learn to distinguish among many forms of data and factors that influence their interpretation in different contexts and over time.
2. Students will learn professional criteria for managing, selecting, and appraising data.
3. Students will learn to use and assess data collections, repositories, and services.
4. Students will gain technical skills in managing data in specific research settings.
5. Students will gain a basic knowledge of practices to curate digital data.
6. Students will learn basic principles of public policies for data.

Course Materials

All course materials will be posted on or linked from the CCLE site for this course. Enrolled students have access to the site at <http://www.ccle.ucla.edu>.

We will refer also to the two books required for 262A: (Borgman, 2015; Ray, 2014)

Office Hours

Wednesdays, 1:30-3:30pm (link posted on CCLE); other times by appointment on Tuesdays and Thursdays, as the instructor is traveling intermittently this term. Office hour meetings can be done by video (Skype, FaceTime), if arranged in advance. Send contact request to CLBORGMAN for Skype or CLBORGMAN@ME.COM for FaceTime.

Grading

Assignment 1 (individual work) 30%
Term project (team work) 50%
Class participation and analysis of readings 20%

Details of the assignments are provided on separate documents.

Students are expected to complete all assigned readings prior to each week's class sessions and come prepared to discuss them. Your preparation and contributions to the discussion are the basis for 20% of your grade. Written assignments are to be submitted electronically to the CCLE site and on paper at the beginning of the class session, as noted. Assignments will be marked down 2 points for each day late. No assignments will be accepted after midnight Monday, June 8.

Summary of Assignment due dates

April 1:	Assignment 1 given
	Term project assignment given
April 15:	Term project proposal due
April 23:	Assignment 1 due
May 11:	Team Project Report outline due
May:	Teams to meet with instructor during office hours
June 3:	Student presentations
June 8 (Monday):	Final projects due to CCLE

Topics, Readings, and Guest Speakers

Readings are to be completed in advance of each class session. Please come to class prepared to discuss the material and its relationship to larger issues in the course. Prepare some talking points as part of your reading and studying.

April 10: This is the finalized order of topics with all guest speakers confirmed

Week 1, April 1: Selection and appraisal

How do we determine what data are worth keeping? What are the criteria? To what extent are these scholarly, disciplinary, policy, or professional questions? Who decides? What role do scholars play in determining what is kept and what are professional roles?

(Borgman, 2015), Chapter 10, What to Keep and Why
(Faniel & Jacobsen, 2010) How scientists assess data for reuse
(Gutmann et al., 2009) ICPSR and Data-PASS on data preservation
(Harvey, 2008) Appraisal and selection from DCC
(A. A. Goodman et al., 2014) Advice to scientists for what to keep and how
(National Health and Medical Research Council, 2007) Australian code for responsible data management
(Pfeiffenberger & Carlson, 2011) Selecting data for publishing in a journal
(“Strategic Directions: Appraisal Policy,” 2007) US government policy on appraisal of records

Week 2, April 8: Data management in the field/data management in the university
Peter Darch and Elizabeth Stephenson

Data are managed during the research process in much different ways than they are managed later, when contributed to data repositories, libraries, or archives. These

Data management in the field: Peter Darch

Researchers are responsible for managing their own data when in the field. Multiple factors can complicate management: data collection plans change dynamically, adapting to unanticipated challenges or new possibilities that emerge; equipment can fail, or prove unsuitable in physical conditions, and limited time means scientists prioritize data collection at the expense of recordkeeping and other bureaucratic tasks.

(Barkhuus & Brown, 2012) Social science fieldwork and collaboration

(Bird, Willoughby, & Frey, 2013) Lab notebooks in the digital era

(Borgman, Wallis, & Enyedy, 2007) Field research in ecology, and how they manage their data

(Lindseth & Baker, 2012) Designing a data logger for oceanography

Data Management in the University: Elizabeth Stephenson

Once scholars are ready to release their data, they may offer them to a repository, which may or may not accept those datasets. Data collections and repositories take many forms and reside in many places. Libbie Stephenson will provide an introduction to data repositories, to the technical aspects of data ingest, and professional aspects of repository management. Read the two journal articles on repository management and browse the repository links.

Repositories – Management issues

[\(Kimpton & Morris, 2013; Peer, Green, & Stephenson, 2014\)](#)

Repositories - offering services

Australian National Data Service. (2014). Retrieved January 24, 2014, from

<http://www.ands.org.au/>

UK Data Service <http://ukdataservice.ac.uk/>

Inter-university Consortium for Political and Social Research (ICPSR)

<http://www.icpsr.umich.edu>

DataBib <http://databib.org/>

Registry of Research Data Repositories <http://www.re3data.org/>

OpenContext <http://opencontext.org/> and The Digital Archaeological record (tDAR)

<https://www.tdar.org/>

Repositories - self-deposit

Dryad <http://datadryad.org/>

figshare <http://figshare.com/>

DataShare/U.S. <http://datashare.ucsf.edu/xtf/search>

DataShare/U.K. <http://datashare.is.ed.ac.uk/>
openICPSR <http://openicpsr.org/>

Repository Systems

Islandora <http://islandora.ca/>
Duraspace <http://www.duraspace.org/>
Hydra Project <http://projecthydra.org/>
California Digital Library Merritt <http://www.cdlib.org/uc3/merritt/>

Week 3, April 15: Technologies of persistence and identification

The technologies associated with data curation and management vary widely by domain, method, and many other factors. These technical overviews will aid you in developing your term project and doing assignment 1. We will focus on generic tools and principles for technology as an entry point for developing a broader background. These readings and the video address the following topics:

(“Data Foundation and Terminology (DFT) WG Recommendations,” 2015) RDA; basic terminology and models for data handling

Persistent identification
Unique identification of digital objects
Digital Object Identifiers
CrossRef
Open Archives Initiative protocols OAI-PMH and OAI-ORE
Research Objects and provenance
Linked open data for data
Namespaces, URLs, and versions of record

(“Data Foundation and Terminology (DFT) WG Recommendations,” 2015) RDA; basic terminology and models for data handling

Research objects:

(Bechhofer et al., 2013; Bechhofer, De Roure, Gamble, Goble, & Buchan, 2010; CrossRef, 2009; A. Goodman et al., 2014; “OpenURL and CrossRef,” 2006; Pepe, Mayernik, Borgman, & Van de Sompel, 2010; Sanderson & Van de Sompel, 2012; Simons, 2012; Van de Sompel, 2013; Weigel, Lautenschlager, Toussaint, & Kindermann, 2013)

Recommended:

(Pepe, Mayernik, Borgman, & Van de Sompel, 2010) Object Reuse and Exchange; life cycles of data objects

Week 4, April 22: Data, Publishing, and Publishers: 8am-11:20am session

This class session will start at 8am for a live broadcast of Prof. Borgman's keynote presentation to the Association of Science, Technology, and Medicine Publishers in Washington, DC, followed by a panel discussion. Dr. Darch will host the session at UCLA and lead the remainder of the discussion locally.

Note: These readings changed from the April 1 version of the syllabus

Required readings:

(Bourne et al., 2011) Force11 Manifesto on scholarly communication and publishing

(Borgman, 2015, Chapter 9) Credit, attribution, and discovery

(Starr et al., 2015) Data citation for accessibility

("Brussels Declaration," 2007) Publishers' statement on data and publication

(Parsons & Fox, 2013) Is data publication the right metaphor?

(Lagoze, Edwards, Sandvig, & Plantin, 2015) Alternative infrastructures in Scholarly publishing

Sites to browse for background:

("RDA/WDS Publishing Data IG," 2015) see summary at
<https://docs.google.com/folderview?usp=sharing&id=0B4qnUFYMGSc-eGFJaGR4NXJpVVk>

(DataCite, 2013) Consortium

(Thomson Reuters, 2015) Commercial index to data

(Society for Scholarly Publishing, 2015) Scholarly Kitchen Blog by SSP

(Macmillan, 2015) Digital Science, aggregation of multiple ventures from the publisher of Nature

(Howard, 2013; "LIBER response to Elsevier's text and data mining policy | LIBER," 2014, "The Cost of Knowledge," 2012; Shillum, 2014) Multiple sources of Elsevier policies; boycott petition

(Force11, 2015) new models of scholarly communication

Recommended:

(Borgman, 2003) The invisible library
(Edwards et al., 2013) Knowledge infrastructures
(Ginsparg, 2011) history of ArXiv

Week 5, April 29: Economics of data: preservation, access, and sustainability

In managing and curating research data, the economic issues are unavoidable. What are the economic models for sustaining access to data? How can projects be costed and budgeted? How do economic considerations influence choices of what to keep, for how long, and why? We will consider several approaches to the economics of research data.

Note: please read ahead for the Provenance topic. Today we will finalize questions to ask Dr. Groth next week, so we can send them to him in advance.

(Beagrie & Houghton, 2014) Value of data sharing and curation
(Berman et al., 2010) US policy study on economics of digital preservation
(David, 2009) Commons and anti-commons for access to research information
(Hess & Ostrom, 2007) Overview of commons models of information
(Kelty, 2012) open access, open science, and motivations of scholars; free riders
(Rosenthal, 2010) Cost analysis for digital preservation of data; keynote presented as long blog post

Week 6, May 6: Provenance in data: WWW and archival approaches, Paul Groth, Free University of Amsterdam and Elsevier Research, by video

Archival and Internet approaches, Paul Groth, University of Amsterdam, discussant, 9am-10:30am by video. Groth will discuss which of the many competing definitions of provenance is most useful for data, what aspects of provenance are most applicable to data management, to curation, to discovery, and to curation, what are the points of intersection between archival, historical, and technical concepts of provenance, especially with respect to data, and what are the practical considerations in documenting provenance for data stewardship.

Note: please read ahead for the Intellectual Property topic. Today we will finalize questions to ask Peter Hirtle, so we can send them to him in advance.

Note: Please read ahead for the options for Week 9 topics so that we can finalize our choice and prepare accordingly.

Readings:

(Barbier, Feng, Gundecha, & Liu, 2013; Bunn, 2014; Cheny, Gil, Groth, & Miles, 2011; Garber, 2012; A. A. Goodman et al., 2014; Groth & Moreau, 2013; Meng, 2011; Niu, 2013)

Week 7, May 13: Ethics in use of public data for research Prof Katie Shilton, U of Maryland, and Jeffrey Burke, UCLA

Prof. Shilton, a PhD graduate of the UCLA Information Studies program, and Prof. Jeffrey Burke of the UCLA Design and Media Arts, are conducted NSF-funded research on Named Data Networking and policy impacts thereof. They focus today on research ethics and open data.

(Boyd & Crawford, 2012) Critical questions about the use of data in research
(Kramer, Guillory, & Hancock, 2014) Notorious Facebook study and responses in media
(Goel, 2014a, 2014b; Sullivan, 2014)
(Shilton, 2012) Participatory research data in the information sciences
(Zimmer, 2010). Ethics of research on Facebook

Week 8, May 20: Intellectual property in data: Peter Hirtle, Berkman Center, Harvard, by video

Policies and practices for open data tend to say little about the problems of intellectual property rights. The OECD report (read earlier in the term) is an exception. Yet rights in data are unclear and often contested. Investigators may not know what rights they have in data, and they may be obtaining data from other sources for which they have little control over rights, formats, reuse, deposit, etc. Complicating matters further are the license stacking problems, wherein each dataset has particular rights associated with it, making data difficult to combine.

(Australian Law Reform Commission, 2014; Ball, 2012; Hirtle, 2011; “LIBER response to Elsevier’s text and data mining policy | LIBER,” 2014; Scotchmer, 2003; Shillum, 2014) Elsevier text mining policy

Week 9: choice of two topics:

Week 9, May 25: University-Industry partnerships: Data transfer

University-industry partnerships that involve the transfer of technologies developed in university research for commercialization are increasingly common. Central to the success of these transfers is successful exchange of data used in the development and testing of these technologies. This exchange typically involves negotiations, which can include: issues of intellectual property and ownership, protection of research subjects; and data standards and formats.

Reading:

(Evans, 2010)

(Mowery, 2004) – Chapter 5 (pdf available online)

(UCLA Office of Intellectual Property and Industry Sponsored Research, 2014)

(Louis, Jones, Anderson, Blumenthal, & Campbell, 2001)

Two short videos:

(*UCLA Technology Transfer and Startups*, 2014),

<https://www.youtube.com/watch?v=Idun5eYjvS4>

(*UCLA Inventor*, 2014),

https://www.youtube.com/watch?v=nZymuXOrC5Q&feature=youtube_gdata_player

Week 9, May 25: Data in public engagement

Libraries, archives, and academic researchers face increasing pressure to engage diverse audiences beyond the university. They may make special collections accessible either physically or virtually, or recruit members of the public to participate in citizen science projects. These processes of public engagement involve selection of research data or artefacts, transforming these research objects accessible into more accessible objects (such as visualizations), and training the public in handling, processing, or interpreting these objects.

Reading:

(Darch, 2014)

(Fortson et al., 2011)

(Raddick et al., 2013)

(Causer & Wallace, 2012)

(Lang & Rio-Ross, 2011)

(Prestopnik & Crowston, 2011)

Week 10, June 3: Project presentations

See project assignment for details. We will devote the last class session to a public presentation of student projects and to a general discussion of the data curation issues identified in each project.

Week 11 (Exam Week), June 8 (Monday): Projects due

Final projects are due by midnight on Monday, June 8.

References

- Australian Law Reform Commission. (2014). Defining “non-consumptive” use. Retrieved January 15, 2014, from <http://www.alrc.gov.au/publications/8-non-consumptive-use/defining-%E2%80%98non-consumptive%E2%80%99-use>
- Ball, A. (2012). How to License Research Data. Retrieved March 3, 2013, from <http://www.dcc.ac.uk/resources/how-guides/license-research-data>
- Barbier, G., Feng, Z., Gundecha, P., & Liu, H. (2013). Provenance Data in Social Media. *Synthesis Lectures on Data Mining and Knowledge Discovery*, 4(1), 1–84. <http://doi.org/10.2200/S00496ED1V01Y201304DMK007>
- Barkhuus, L., & Brown, B. (2012). The Sociality of Fieldwork: Designing for Social Science Research Practice and Collaboration. In *Proceedings of the 17th ACM International Conference on Supporting Group Work* (pp. 35–44). New York, NY, USA: ACM. <http://doi.org/10.1145/2389176.2389183>
- Beagrie, N., & Houghton, J. (2014). The Value and Impact of Data Sharing and Curation. JISC. Retrieved from repository.jisc.ac.uk/5568/1/iDF308_-_Digital_Infrastructure_Directions_Report%2C_Jan14_v1-04.pdf
- Bechhofer, S., Buchan, I., De Roure, D., Missier, P., Ainsworth, J., Bhagat, J., ... Goble, C. (2013). Why linked data is not enough for scientists. *Future Generation Computer Systems*, 29(2), 599–611. <http://doi.org/10.1016/j.future.2011.08.004>
- Bechhofer, S., De Roure, D., Gamble, M., Goble, C., & Buchan, I. (2010). Research Objects: Towards Exchange and Reuse of Digital Knowledge. *Nature Precedings*. <http://doi.org/10.1038/npre.2010.4626.1>
- Berman, F., Lavoie, B., Ayris, P., Choudhury, G. S., Cohen, E., Courant, P., ... Van Camp, A. (2010). *Sustainable Economics for a Digital Planet: Ensuring Long-Term Access to Digital Information*. San Diego: National Science Foundation, Andrew W. Mellon Foundation, Library of Congress, Joint Information Systems Committee of the UK, Council on Library and Information Resources, National Archives and Records Administration. Retrieved from <http://brtf.sdsc.edu/publications.html>
- Bird, C. L., Willoughby, C., & Frey, J. G. (2013). Laboratory notebooks in the digital era: the role of ELNs in record keeping for chemistry and other sciences. *Chemical Society Reviews*, 42(20), 8157. <http://doi.org/10.1039/c3cs60122f>
- Borgman, C. L. (2003). The Invisible Library: Paradox of the Global Information Infrastructure. *Library Trends*, 51, 652–674.
- Borgman, C. L. (2015). *Big Data, Little Data, No Data: Scholarship in the Networked World*. Cambridge MA: MIT Press.
- Borgman, C. L., Wallis, J. C., & Enyedy, N. D. (2007). Little Science Confronts the Data Deluge: Habitat Ecology, Embedded Sensor Networks, and Digital Libraries. *International Journal on Digital Libraries*, 7(1-2), 17–30. <http://doi.org/10.1007/s00799-007-0022-9>
- Bourne, P. E., Clark, T., Dale, R., de Waard, A., Hovy, E. H., & Shotton, D. (Eds.). (2011). Force 11 Manifesto: Improving Future Research Communication and e-Scholarship. Retrieved from http://www.force11.org/white_paper

- Boyd, danah, & Crawford, K. (2012). Critical Questions for Big Data. *Information, Communication & Society*, 15(5), 662–679.
<http://doi.org/10.1080/1369118X.2012.678878>
- Brussels Declaration. (2007). Retrieved from <http://www.stm-assoc.org/public-affairs/resources/brussels-declaration/>
- Bunn, J. (2014). Questioning autonomy: an alternative perspective on the principles which govern archival description. *Archival Science*, 14(1), 3–15.
<http://doi.org/10.1007/s10502-013-9200-2>
- Causer, T., & Wallace, V. (2012). Building a volunteer community: results and findings from Transcribe Bentham. *Digital Humanities Quarterly*, 6. Retrieved from <http://discovery.ucl.ac.uk/1362050/>
- Chen, J., Gil, Y., Groth, P., & Miles, S. (2011). *Requirements for Provenance on the Web*. Retrieved from http://www.w3.org/2005/Incubator/prov/wiki/User_Requirements
- CrossRef. (2009). *The Formation of CrossRef: A Short History*. CrossRef. Retrieved from <http://www.crossref.org/01company/02history.html>
- Darch, P. T. (2014, February). *Managing the Public to Manage Data: Citizen Science and Astronomy*. Presented at the 9th International Digital Curation Conference, San Francisco, CA. Retrieved from <http://www.dcc.ac.uk/events/idcc14/programme-presentations#sthash.UWsfM9QB.dpuf>
- DataCite. (2013). [Home page]. Retrieved September 10, 2013, from <http://www.datacite.org/>
- Data Foundation and Terminology (DFT) WG Recommendations. (2015). Retrieved March 31, 2015, from <https://rd-alliance.org/group/data-foundation-and-terminology-wg/outcomes/data-foundation-and-terminology-dft-wg>
- David, P. (2009). *Mitigating “Anticommons” Harms to Research In Science and Technology* (Discussion Paper No. 10-009). Stanford Institute for Economic Policy Research. Retrieved from <http://ideas.repec.org/p/sip/dpaper/10-009.html>
- Edwards, P. N., Jackson, S. J., Chalmers, M. K., Bowker, G. C., Borgman, C. L., Ribes, D., ... Calvert, S. (2013). *Knowledge Infrastructures: Intellectual Frameworks and Research Challenges*. Ann Arbor: University of Michigan. Retrieved from <http://deepblue.lib.umich.edu/handle/2027.42/97552>
- Evans, J. A. (2010). Industry collaboration, scientific sharing and the dissemination of knowledge. *Social Studies of Science*. Retrieved from <http://sss.sagepub.com/content/early/2010/09/02/0306312710379931.abstract>
- Faniel, I. M., & Jacobsen, T. E. (2010). Reusing Scientific Data: How Earthquake Engineering Researchers Assess the Reusability of Colleagues’ Data. *Journal of Computer-Supported Cooperative Work*, 19, 355–375.
<http://doi.org/10.1007/s10606-010-9117-8>
- Force11. (2015). Home page. Retrieved August 6, 2014, from <https://www.force11.org/about>
- Fortson, L., Masters, K., Nichol, R., Borne, K. D., Edmondson, E., Lintott, C., ... Wallin, J. (2011). Galaxy Zoo: Morphological Classification and Citizen Science. *arXiv:1104.5513*. Retrieved from <http://arxiv.org/abs/1104.5513>

- Garber, M. (2012, March 15). Scholars: Yes, We Need Better Attribution Systems (but No, We Don't Know How to Make Them, Either). Retrieved March 31, 2014, from <http://www.theatlantic.com/technology/archive/2012/03/scholars-yes-we-need-better-attribution-systems-but-no-we-dont-know-how-to-make-them-either/254527/>
- Ginsparg, P. (2011). ArXiv at 20. *Nature*, 476(7359), 145–147. <http://doi.org/10.1038/476145a>
- Goel, V. (2014a, June 29). Facebook Tinkers With Users' Emotions in News Feed Experiment, Stirring Outcry. *The New York Times*. Retrieved from <http://www.nytimes.com/2014/06/30/technology/facebook-tinkers-with-users-emotions-in-news-feed-experiment-stirring-outcry.html>
- Goel, V. (2014b, August 12). As Data Overflows Online, Researchers Grapple With Ethics. *The New York Times*. Retrieved from <http://www.nytimes.com/2014/08/13/technology/the-boon-of-online-data-puts-social-science-in-a-quandary.html>
- Goodman, A. A., Pepe, A., Blocker, A. W., Borgman, C. L., Cranmer, K., Crosas, M., ... Slavkovic, A. (2014). 10 Simple Rules for the Care and Feeding of Scientific Data. *PLoS Computational Biology*. Retrieved from https://www.authorea.com/users/3/articles/3410/_show_article
- Goodman, A., Pepe, A., Blocker, A. W., Borgman, C. L., Cranmer, K., Crosas, M., ... Slavkovic, A. (2014). Ten Simple Rules for the Care and Feeding of Scientific Data. *PLoS Computational Biology*, 10(4), e1003542. <http://doi.org/10.1371/journal.pcbi.1003542>
- Groth, P., & Moreau, L. (2013). PROV-Overview. Retrieved April 14, 2014, from <http://www.w3.org/TR/2013/NOTE-prov-overview-20130430/>
- Gutmann, M., Abrahamson, M., Adams, M., Altman, M., Arms, C. R., & King, G. (2009). From Preserving the Past to Preserving the Future: The Data-PASS Project and the Challenges of Preserving Digital Social Science Data. *Library Trends*, 57, 315–337. <http://doi.org/10.1353/lib.0.0039>
- Harvey, R. (2008). Appraisal and Selection. Digital Curation Centre. Retrieved from <http://www.dcc.ac.uk/resources/briefing-papers/introduction-curation/appraisal-and-selection>
- Hess, C., & Ostrom, E. (2007). Introduction: An overview of the knowledge commons. In C. Hess & E. Ostrom (Eds.), *Understanding knowledge as a commons: from theory to practice* (pp. 3–26). Cambridge, MA: MIT Press.
- Hirtle, P. B. (2011). Introduction to Intellectual Property Rights in Data Management. Retrieved March 3, 2013, from <https://confluence.cornell.edu/display/rdmsgweb/introduction-intellectual-property-rights-data-management>
- Howard, J. (2013, December 6). Posting Your Latest Article? You Might Have to Take It Down. Retrieved from <http://chronicle.com/blogs/wiredcampus/posting-your-latest-article-you-might-have-to-take-it-down/48865>
- Kelty, C. M. (2012). This is not an article: Model organism newsletters and the question of “Open science.” *BioSocieties*, 7(2), 140–168. <http://doi.org/10.1057/biosoc.2012.8>

- Kimpton, M., & Morris, C. M. (2013). Managing and Archiving Research Data: Local Repository and Cloud-Based Practices. In J. M. Ray (Ed.), *Research Data Management: Practical Strategies for Information Professionals* (pp. 223–238). West Lafayette, IL: Purdue University Press.
- Kramer, A. D. I., Guillory, J. E., & Hancock. (2014). Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences*, 111(24), 8788–8790.
<http://doi.org/10.1073/pnas.1320040111>
- Lagoze, C., Edwards, P., Sandvig, C., & Plantin, J.-C. (2015). Should I Stay or Should I Go? Alternative Infrastructures in Scholarly Publishing. *International Journal of Communication*, 9(0), 20. Retrieved from
<http://ijoc.org/index.php/ijoc/article/view/2929>
- Lang, A., & Rio-Ross, J. (2011). Using Amazon Mechanical Turk to transcribe historical handwritten documents. *The Code4Lib Journal*. Retrieved from
[http://journal.code4lib.org/articles/6004?utm_source=feedburner&utm_medium=feed&utm_campaign=Feed%3A+c4lj+\(The+Code4Lib+Journal\)](http://journal.code4lib.org/articles/6004?utm_source=feedburner&utm_medium=feed&utm_campaign=Feed%3A+c4lj+(The+Code4Lib+Journal))
- LIBER response to Elsevier's text and data mining policy | LIBER. (2014, March 28). Retrieved May 12, 2014, from <http://www.libereurope.eu/news/liber-response-to-elsevier%E2%80%99s-text-and-data-mining-policy>
- Lindseth, B., & Baker, K. (2012). Collaborative Design of an Oceanographic Event Logger. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work* (pp. 1195–1198). New York, NY, USA: ACM.
<http://doi.org/10.1145/2145204.2145383>
- Louis, K. S., Jones, L. M., Anderson, M. S., Blumenthal, D., & Campbell, E. G. (2001). Entrepreneurship, secrecy, and productivity: a comparison of clinical and non-clinical life sciences faculty. *The Journal of Technology Transfer*, 26(3), 233–245.
- Macmillan. (2015). Home. Retrieved April 10, 2015, from <http://www.digital-science.com/>
- Meng, X.-L. (2011). Multi-party inference and uncongeniality. In M. Lovric (Ed.), *International Encyclopedia of Statistical Science* (pp. 884–888). Berlin: Springer.
- Mowery, D. C. (2004). *Ivory tower and industrial innovation: university-industry technology transfer before and after the Bayh-Dole act in the United States*. Stanford University Press. Retrieved from
https://books.google.com/books?hl=en&lr=&id=sCscGlm2Q8YC&oi=fnd&pg=PR13&dq=%22what+happens+in+university-industry+technology+transfer%22+lowe+mowery&ots=cbLpRT3uFR&sig=b-Y_EQnGxtclCsX39w802PcjIII
- National Health and Medical Research Council. (2007). Australian Code for the Responsible Conduct of Research. Retrieved January 25, 2014, from
<http://www.nhmrc.gov.au/guidelines/publications/r39>
- Niu, J. (2013). Provenance: crossing boundaries. *Archives and Manuscripts*, 41(2), 105–115. <http://doi.org/10.1080/01576895.2013.811426>
- OpenURL and CrossRef. (2006). Retrieved from
<http://www.crossref.org/02publishers/16openurl.html>

- Parsons, M. A., & Fox, P. A. (2013). Is data publication the right metaphor? *Data Science Journal*, 12, WDS32–WDS46. <http://doi.org/10.2481/dsj.WDS-042>
- Peer, L., Green, A., & Stephenson, E. (2014). Committing to Data Quality Review. *International Journal of Digital Curation*, 9(1), 263–291. <http://doi.org/10.2218/ijdc.v9i1.317>
- 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, 567–582. <http://doi.org/10.1002/asi.21263>
- Pfeiffenberger, H., & Carlson, D. (2011). “Earth System Science Data” (ESSD) - A Peer Reviewed Journal for Publication of Data. *D-Lib Magazine*, 17. <http://doi.org/10.1045/january2011-pfeiffenberger>
- Prestopnik, N. R., & Crowston, K. (2011). Gaming for (citizen) science: exploring motivation and data quality in the context of crowdsourced science through the design and evaluation of a social-computational system. In *e-Science Workshops (eScienceW), 2011 IEEE Seventh International Conference on* (pp. 28–33). IEEE. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6130727
- Raddick, M. J., Bracey, G., Gay, P. L., Lintott, C. J., Cardamone, C., Murray, P., ... Vandenberg, J. (2013). Galaxy Zoo: Motivations of citizen scientists. *Astronomy Education Review*, 12(1), 010106.
- Ray, J. M. (2014). *Research data management: practical strategies for information professionals*. West Lafayette, Ind.: Purdue University Press.
- Rosenthal, D. S. H. (2010). Stepping Twice Into The Same River. Presented at the Joint Conference on Digital Libraries, New York: ACM. Retrieved from <http://blog.dshr.org/2010/06/jcdl-2010-keynote.html>
- Sanderson, R., & Van de Sompel, H. (2012). Cool URIs and dynamic data. *IEEE Internet Computing*, 16(4), 76–79. <http://doi.org/10.1109/MIC.2012.78>
- Scotchmer, S. (2003). Intellectual property—When is it the best incentive mechanism for S&T data and information? In J. M. Esanu & P. F. Uhler (Eds.), *The Role of Scientific and Technical Data and Information in the Public Domain* (pp. 15–18). Washington, DC: National Academies Press.
- Shillum, C. (2014, January 31). Elsevier updates text-mining policy to improve access for researchers. Retrieved May 12, 2014, from <http://www.elsevier.com/connect/elsevier-updates-text-mining-policy-to-improve-access-for-researchers>
- Shilton, K. (2012). Participatory personal data: An emerging research challenge for the information sciences. *Journal of the American Society for Information Science and Technology*, 63(10), 1905–1915. <http://doi.org/10.1002/asi.22655>
- Simons, N. (2012). Implementing DOIs for research data. *D-Lib Magazine*, 18(5/6). <http://doi.org/10.1045/may2012-simons>
- Society for Scholarly Publishing. (2015). The Scholarly Kitchen. Retrieved April 10, 2015, from <http://scholarlykitchen.sspnet.org/>
- Starr, J., Castro, E., Crosas, M., Dumontier, M., Downs, R. R., Duerr, R., ... Clark, T. (2015). *Achieving human and machine accessibility of cited data in scholarly publications* (No. e697v4). PeerJ PrePrints. Retrieved from <https://peerj.com/preprints/697v4>

- Strategic Directions: Appraisal Policy. (2007). Retrieved March 29, 2014, from <http://www.archives.gov/records-mgmt/initiatives/appraisal.html>
- Sullivan, G. (2014, July 1). Cornell ethics board did not pre-approve Facebook mood manipulation study. *The Washington Post*. Retrieved from <http://www.washingtonpost.com/news/morning-mix/wp/2014/07/01/facebooks-emotional-manipulation-study-was-even-worse-than-you-thought/>
- The Cost of Knowledge. (2012). Retrieved April 10, 2015, from <http://thecostofknowledge.com/>
- Thomson Reuters. (2015). Data Citation Index. Retrieved October 18, 2012, from http://wokinfo.com/products_tools/multidisciplinary/dci/
- UCLA Inventor: Dr. Patricia Ganz*. (2014). Retrieved from https://www.youtube.com/watch?v=nZymuXOrC5Q&feature=youtube_gdata_player
- UCLA Office of Intellectual Property and Industry Sponsored Research. (2014). *FAQ and Guide: Technology Transfer*. Retrieved from <http://www.research.ucla.edu/tech/techtransferFAQ.pdf>
- UCLA Technology Transfer and Startups*. (2014). Retrieved from https://www.youtube.com/watch?v=Idun5eYjvS4&feature=youtube_gdata_player
- Van de Sompel, H. (2013, April). *From the Version of Record to a Version of the Record*. Presented at the Coalition for Networked Information. Retrieved from http://www.youtube.com/watch?v=fhrGS-QbNVA&feature=youtube_gdata_player
- Weigel, T., Lautenschlager, M., Toussaint, F., & Kindermann, S. (2013). A Framework for Extended Persistent Identification of Scientific Assets. *Data Science Journal*, 12, 10–22. <http://doi.org/10.2481/dsj.12-036>
- Zimmer, M. (2010). “But the data is already public”: on the ethics of research in Facebook. *Ethics and Information Technology*, 12(4), 313–325. <http://doi.org/10.1007/s10676-010-9227-5>