

UCLA

Publications

Title

Syllabus for Data Management and Practice, Part I, Winter 2015, UCLA Information Studies

Permalink

<https://escholarship.org/uc/item/959619q4>

Author

Borgman, Christine L.

Publication Date

2015

Data Management and Practice

Winter, 2015, UCLA Information Studies 262A,
Mon 1:30-4:50, IS Room 121, January 5 through March 16 (exam week)
Christine L. Borgman, Professor & Presidential Chair
215 GSE&IS bldg, 310-825-6164; Christine.Borgman@ucla.edu
[Spring 2015, 262B: Weds 9-12:20]

Table of Contents

Course Description: 262A, 262B	2
Course Objectives	2
Course Materials	2
Office Hours.....	3
Grading	3
Summary of Assignment due dates.....	3
Topics, Readings, and Guest Speakers	4
Week 1: Overview of Data Management and Practices, January 5	4
Week 2: What are data? January 12.....	4
Week 3: Martin Luther King Holiday, January 19	5
Week 4: Public policies for research data, requirements for researchers, January 26.....	5
Week 5: Data management plans and processes, February 2	6
Week 6: Data practices in the scientific domains, February 9.....	6
Week 7, President’s holiday, February 16	6
Week 8: Data practices in the social scientific and humanities domains, February 23	7
Week 9: Data sharing and reuse: Practice and policy, March 2	7
Week 10: Data Management by research teams, libraries, and archives, March 9	8
Week 11 (Exam Week): Student presentations, March 16.....	8
Tentative spring term topics.....	9
Syllabus 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

Two books are required for purchase: (Borgman, 2015a; Ray, 2014)

Note: The Borgman book will be available in the UCLA bookstore on or about January 12. Printing will be completed on Jan 2; MIT Press ships advance orders on January 9. Reading assignments are scheduled accordingly.

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

Office Hours

Mondays and Wednesdays, 5-6pm, starting January 7 (link posted on CCLE); other times by appointment and by email.

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 / Moodle 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 1:30pm on Monday, March 16. No exceptions, as the instructor is leaving the country on March 19.

Summary of Assignment due dates

January 21 (Weds): Assignment 1, part 1
January 22 (Thurs): Term project proposal
February 2 (Mon): Assignment 1, full report
February 19 (Thurs): Team Project Report outline
February 26 (Thurs): Daniel Rosenberg colloquim (recommended)
Feb 2-Mar 9: Teams to meet with instructor during office hours
March 16 (Mon, Exam week): Student presentations and final projects due

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.

Week 1: Overview of Data Management and Practices, January 5

We will devote the first week of class to an overview of the concepts of data, management, and applications across scholarly disciplines. We will begin to form project groups for the term. Assignment #1 and the term project will be distributed and explained.

Copies of Borgman 2015, Chaps 1 and 2, will be distributed in class, in advance of the January 9 shipping date of the book.

Readings

(Ray, 2014), Introduction, pages 1-21

(Ayres, 2007), Introduction, p 1-18

(Cukier & Mayer-Schoenberger, 2013) Overview of their book (Mayer-Schoenberger & Cukier, 2013)

(Lazer, Kennedy, King, & Vespignani, 2014) Google Flu

(Rosenberg, 2003) History of information overload

Video (4:49): <http://www.youtube.com/watch?v=j50ZssEojtM>

Week 2: What are data? January 12

“Data” is a far more ambiguous concept than is immediately apparent. Decisions about what data are to be managed, shared, and curated depend heavily on how the concept is defined. We will explore definitions and facets of “data” as a basis for discussion throughout the term. Project teams and assignments will be made today.

Assignment: Bring in a sample today of something that you consider to be data. We will discuss them in class.

Readings

(Borgman, 2015a), Chapters 1 and 2, Provocations; What Are Data?

(Edwards et al., 2013); see also Knowledge Infrastructures site:

<http://knowledgeinfrastructures.org>

(Rosenberg, 2013) History of “data”

(Laney, 2001) Volume, variety, velocity

Week 3: Martin Luther King Holiday, January 19

Teams are to have organizational meetings this week and begin meeting with their UCLA research team partners. Work on assignment 1; bring questions to class next week.

The preliminary part of Assignment 1 is due on Weds, January 21, to CCLE.

Term project proposal is due Thursday, January 22, to CCLE. Please make office hour appointments to discuss your project in the weeks ahead.

Readings:

(Borgman, 2015a) Chapters 1-4

Week 4: Public policies for research data, requirements for researchers, January 26

Researchers' rights and responsibilities for data management are codified in public policies. These policies have legal and economic aspects that vary widely; many international agreements also are in place.

Readings

(Ray, 2014) part 1: Understanding the policy context (2 chapters)

(Steinhart, Chen, Arguillas, Dietrich, & Kramer, 2012) Researchers' responses to data management plan requirements

(Organization for Economic Cooperation and Development, 2007) International policy on data sharing

(Wood et al., 2010) EU Riding the Wave report

(National Science Foundation, 2011a, 2011b) NSF policies on data sharing

(Australian National Data Service, 2014; National Health and Medical Research Council, 2007) Australian policies on data management

(Fox & Harris, 2013) International policy analysis for science

Week 5: Data management plans and processes, February 2

“Data management” encompasses activities performed throughout a research project and well beyond, and can refer to specific plans that are mandated by funding agencies. Some activities are local and ad hoc; some are distributed and standardized; most fall somewhere in between. Researchers, data scientists, data librarians, repository staff, publishers, and many other stakeholders may be involved in data management. This week we address basic principles and components of the planning process.

Assignment #1 is due at the start of class today.

Readings

(Ray, 2014), Part 2, Planning for data management, chapters 3-5
 (“UC3: Data Management Guidelines,” 2014) California Digital Library
 (“European Landscape Study of Research Data Management,” 2013) European study of data management needs of researchers
 (Fox & Harris, 2013) Summary of international planning
 (Kimpton & Morris, 2013) Local repository and cloud-based practices (Ray, Ch 11)

Week 6: Data practices in the scientific domains, February 9

Notions of data vary greatly by context, discipline, time, and place. We will spend weeks 6 and 7 exploring case studies in multiple fields. Much research policy and data management practice is based on scientific data, thus we start with the sciences.

Readings

(Borgman, 2015a), Chapter 5, Science cases
 (National Science Board, 2005)
 (Borgman, Wallis, & Mayernik, 2012) Science and computer science practices
 (Ribes & Jackson, 2013) Data in science collaborations
 (Edwards, Mayernik, Batcheller, Bowker, & Borgman, 2011) Data as glue and friction

Week 7, President’s holiday, February 16

Please use this week to work together on your team projects, especially drafting your final report and presentation.

Team Project Report outline due Thursday, February 19.

Week 8: Data practices in the social scientific and humanities domains, February 23

Data management practices in the social sciences and humanities tend to be much different from those in the sciences. UCLA has deep expertise in these areas, hence leaders in these areas are invited for a class discussion this week.

Tentative Panel Session:

Elizabeth Stephenson, Director, [Institute for Social Research Data Archive](#), UCLA

Dr. Lisa Snyder, Institute for Digital Research and Education, UCLA

<https://idre.ucla.edu/people/profiles/lisa-snyder>

Dr. Miriam Posner, Program Coordinator, Digital Humanities Program, UCLA

<http://miriamposner.com/about.html>

Note: Prof. Daniel Rosenberg is speaking in the IS Dept Colloquium this week about his research on the history of data: Thursday, Feb 26, 3-5pm, GSEIS 111.

Readings

(Borgman, 2015a) Chapter 5: Social sciences cases; Chapter 6, Humanities cases

(Ray, 2014), Chapter 10, Social Science Data

(Vardigan & Whiteman, 2007) ICPSR and OAIS

(Inter-university Consortium for Political and Social Research, 2012) Best practices guide for social sciences data

(King, 2011) Stewardship of social sciences data

(“Archaeology Data Service,” 2013) Best practices guide

(Arts & Humanities Research Council, 2012) Technical planning guide

(Borgman, 2009) Challenge to the humanities in uses of data

Week 9: Data sharing and reuse: Practice and policy, March 2

Despite the proliferation of data sharing policies, many factors augur against data sharing and reuse. Incentives to release data often run counter to the reward systems of scholarship; skills and resources are lacking, and suitable repositories may not exist. This week we will discuss practice, policy, and perspectives of the many stakeholders in data sharing and reuse.

Readings

(Borgman, 2015a), Chapter 8: releasing, sharing, and reusing data

(Wessels et al., 2014) Issues in the development of open access to research data

(Ray, 2014), Chapter 19, Clifford Lynch

(Borgman, 2015b) Short essay on issues in data sharing for EU community

Week 10: Data Management by research teams, libraries, and archives, March 9

We will conclude this term and lay the foundation for Part II of this course with a discussion of the workforces and institutional activities associated with managing data.

Readings

(Borgman, 2015a) Chapter 10, What to keep and why

(Ray, 2014), Chapters 7-10, Managing project data; digital repositories

("European Landscape Study of Research Data Management," 2013) SURF study

(Hanson, Surkis, & Yacobucci, 2012) Short clever story about data management by teams

(Fearon, Gunia, Lake, Pralle, & Sallans, 2013) ARL survey of data management planning; read Executive Summary

Week 11 (Exam Week): Student presentations, March 16

See project assignment for details. We will devote the last class session to a public presentation of student projects and a general discussion of project findings. The last 30 minutes of class are reserved for a discussion of spring term topics (see list below).

Final projects due noon, March 16, to CCLE, with paper copies deposited with the instructor at the start of class.

NB: No extensions as Prof. Borgman is leaving the U.S. on March 19.

Tentative spring term topics

Other topics welcomed; nine topics to be selected (Week 10 reserved for project presentations)

Selection and appraisal of data

Data archives and repositories

Economics of data: preservation, access, and sustainability

Provenance in data: WWW and archival approaches, Paul Groth, Free University of Amsterdam / Elsevier Science (by video)

Data Governance: UCLA Joint Administration-Academic Senate Task Force, Kent Wada, UCLA Chief Privacy Officer

U.S. and international data policy: Paul Uhlir, National Academies of Science

Technologies of persistence and identification

Technical and policy infrastructure for data management. Profs. Katie Shilton, University of Maryland, and Jeffrey Burke, UCLA, on National Data Network research.

Intellectual property in data. Peter Hirtle and staff of the Berkman Center for Law and Policy, Harvard University (by video)

Syllabus References

- Archaeology Data Service / Digital Antiquity: Guides to Good Practices. (2013). Retrieved June 11, 2013, from <http://guides.archaeologydataservice.ac.uk/>
- Arts & Humanities Research Council. (2012). Technical Plan. Retrieved July 2, 2013, from <http://www.ahrc.ac.uk/Funding-Opportunities/Research-funding/RFG/Application-guidance/Pages/Technical-Plan.aspx>
- Australian National Data Service. (2014). [Home page]. Retrieved January 24, 2014, from <http://www.and.s.org.au/>
- Ayres, I. (2007). *Super Crunchers: Why thinking-by-numbers is the new way to be smart*. New York: Bantam.
- Borgman, C. L. (2009). The digital future is now: A call to action for the humanities. *Digital Humanities Quarterly*, 3. Retrieved from <http://digitalhumanities.org/dhq/vol/3/4/000077/000077.html>
- Borgman, C. L. (2015a). *Big Data, Little Data, No Data: Scholarship in the Networked World*. Cambridge MA: MIT Press. Retrieved from <http://mitpress.mit.edu/big-data>
- Borgman, C. L. (2015b, January). If data sharing is the answer, what is the question? *ERCIM News*, (100). Retrieved from <http://ercim-news.ercim.eu/>
- Borgman, C. L., Wallis, J. C., & Mayernik, M. S. (2012). Who's Got the Data? Interdependencies in Science and Technology Collaborations. *Computer Supported Cooperative Work*, 21(6), 485–523. doi:10.1007/s10606-012-9169-z
- Cukier, K. N., & Mayer-Schoenberger, V. (2013, May 1). The Rise of Big Data. *Foreign Affairs*, (May/June 2013). Retrieved from <http://www.foreignaffairs.com/articles/139104/kenneth-neil-cukier-and-viktor-mayer-schoenberger/the-rise-of-big-data?page=show#>
- 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>
- Edwards, P. N., Mayernik, M. S., Batcheller, A. L., Bowker, G. C., & Borgman, C. L. (2011). Science Friction: Data, Metadata, and Collaboration. *Social Studies of Science*, 41, 667–690. doi:10.1177/0306312711413314
- European Landscape Study of Research Data Management. (2013). Retrieved September 27, 2013, from <http://www.surf.nl/en/actueel/Pages/EuropeanLandscapeStudyofResearchDataManagement.aspx>
- Fearon, D. J., Gunia, B., Lake, S., Pralle, B. E., & Sallans, A. L. (2013). Research Data Management Services, SPEC Kit 334 (July 2013). Retrieved from <http://publications.arl.org/Research-Data-Management-Services-SPEC-Kit-334/>
- Fox, P., & Harris, R. (2013). ICSU and the Challenges of Data and Information Management for International Science. *Data Science Journal*, 12, WDS1–WDS12. Retrieved from https://www.jstage.jst.go.jp/article/dsj/12/0/12_WDS-001/_article
- Hanson, K., Surkis, A., & Yacobucci, K. (2012). *Data Sharing and Management Snafu in 3 Short Acts*. Retrieved from <http://www.youtube.com/watch?v=N2zK3sAtr-4&sns=em>
- Inter-university Consortium for Political and Social Research. (2012). *Guide to Social Science Data Preparation and Archiving: Best Practice Throughout the Data Life Cycle* (No. 5th

- edition). Ann Arbor, MI: ICPSR. Retrieved from <https://www.icpsr.umich.edu/icpsrweb/content/deposit/guide/index.html>
- 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.
- King, G. (2011). Ensuring the Data-Rich Future of the Social Sciences. *Science*, 331, 719–721. doi:10.1126/science.1197872
- Laney, D. (2001). 3D Data Management: Controlling Data Volume, Velocity, and Variety. Retrieved July 10, 2013, from <http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf>
- Lazer, D., Kennedy, R., King, G., & Vespignani, A. (2014). The Parable of Google Flu: Traps in Big Data Analysis. *Science*, 343(6176), 1203–1205. doi:10.1126/science.1248506
- Mayer-Schonberger, V., & Cukier, K. (2013). *Big Data: A Revolution That Will Transform How We Live, Work, and Think*. Boston: Houghton Mifflin Harcourt.
- 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>
- National Science Board. (2005). *Long-Lived Digital Data Collections*. Retrieved from <http://www.nsf.gov/pubs/2005/nsb0540/>
- National Science Foundation. (2011a). *NSF Data Management Plans*. Washington, DC: NSF. Retrieved from http://www.nsf.gov/pubs/policydocs/pappguide/nsf11001/gpg_2.jsp#dmp
- National Science Foundation. (2011b). *NSF Data Sharing Policy*. Washington, DC: NSF. Retrieved from http://www.nsf.gov/pubs/policydocs/pappguide/nsf11001/aag_6.jsp#VID4
- Organization for Economic Cooperation and Development. (2007). *OECD Principles and Guidelines for Access to Research Data from Public Funding* (pp. 1–24). Retrieved from www.oecd.org/dataoecd/9/61/38500813.pdf
- Ray, J. M. (2014). *Research data management: practical strategies for information professionals*. West Lafayette, Ind.: Purdue University Press.
- Ribes, D., & Jackson, S. J. (2013). Data bite man: The work of sustaining a long-term study. In L. Gitelman (Ed.), *“Raw Data” is an Oxymoron* (pp. 147–166). Cambridge, MA: MIT Press.
- Rosenberg, D. (2003). Early Modern Information Overload. *Journal of the History of Ideas*, 64(1), 1–9. doi:10.1353/jhi.2003.0017
- Rosenberg, D. (2013). Data before the fact. In L. Gitelman (Ed.), *“Raw Data” is an Oxymoron* (pp. 15–40). Cambridge, MA: MIT Press.
- Steinhart, G., Chen, E., Arguillas, F., Dietrich, D., & Kramer, S. (2012). Prepared to Plan? A Snapshot of Researcher Readiness to Address Data Management Planning Requirements. *Journal of eScience Librarianship*, 1(2). doi:10.7191/jeslib.2012.1008
- UC3: Data Management Guidelines. (2014). Retrieved May 14, 2014, from <http://www.cdlib.org/uc3/datamanagement/documenting.html>
- Vardigan, M., & Whiteman, C. (2007). ICPSR Meets OAIS: Applying the OAIS Reference Model to the Social Science Archive Context. Retrieved from <http://deepblue.lib.umich.edu/handle/2027.42/60440>

- Wessels, B., Finn, R. L., Linde, P., Mazzetti, P., Nativi, S., Riley, S., ... Wyatt, S. (2014). Issues in the development of open access to research data. *Prometheus*, 32(1), 49–66.
doi:10.1080/08109028.2014.956505
- 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>