



Vicky Steeves, Visiting Assistant Professor  
Pratt Institute School of Information  
144 W. 14th St., 6<sup>th</sup> Floor  
New York, NY 10011-7301  
Phone: 212-647-7682  
Email: [vsteeves@pratt.edu](mailto:vsteeves@pratt.edu)  
URL: <https://vickysteeves.gitlab.io/lis-628-datalibrarianship>

---

## INFO 628: Data Librarianship and Management Fall 2018

Class Hours: Thursday, 6:30pm – 9:20pm  
Office Hours: By appointment  
Credits: 3  
Prerequisites: None  
Location: PMC 612

### **Bulletin Description:**

The world of data is seemingly a new frontier for libraries, yet in some ways, data and data sets are comparable to other print and electronic resources that librarians historically have been charged with locating, teaching, collecting, organizing, and preserving. This course asks how best we can serve the needs of a burgeoning community of data users/producers while meeting the new challenges that data present to our existing skillsets, workflows, and infrastructure. Topics will include data reference and literacy; archives and repositories; formats and standards; ethics and policy. Statistical/GIS software and research data management are also explored.

### **Detailed Description:**

Class sessions will include lectures, in-class lab activities, and student-led discussions of readings and data-related news. Practitioners in the field will serve as guest lecturers when available and appropriate. The methods, activities, and assignments in this course are designed to (a) maximize peer learning, i.e. opportunities to teach and learn from other students, (b) approximate some of the real world activities and challenges faced by librarians, and (c) get students excited about (rather than intimidated by) this growing niche of librarianship.

### **Course Goals:**

The course provides:

An introduction to concepts and terminology related to data and data services.

Broad overview of the nature and range of data products and producers.

Knowledge of how to develop and provide different tiers of data services (including reference, instruction, and collections development) in a library setting.

Understanding of ethical, social, and political issues related to the creation, use, and reuse of data.

### Student Learning Outcomes:

*By the end of this course, students will be able to:*

Describe forms, formats, and lifecycles of data and how these vary across disciplines.

Practice effective strategies and appropriate sources for locating different kinds of data and statistics.

Construct basic questions and considerations when collecting and appraising data. Self-sufficiently acquire technical knowledge.

Demonstrate the ability to think critically and communicate confidently about issues related to data librarianship.

### Course Schedule and Readings

While this syllabus provides a basic framework for the course, it is subject to change. All changes will be announced in class and on the course website (<https://vickysteeves.gitlab.io/lis-628-datalibrarianship>) and via email. Unless otherwise noted, the readings will be linked below and within each week's module on this site. All readings will be open access, so you will be able to read them without logging into anything. Readings should be completed in advance of the week they're assigned below.

If you see dead links (it does happen, usually with no notice), weird due dates, or other syllabus problems, please email me!

Unit 1: Definitional/technical overview	
Week 1: Course Overview   August 29th	
No readings	Bring in a sample of something that you consider to be data. We will discuss them in class.
Week 2: Data Basics   September 5th	
Borgman, " <a href="#">The conundrum of sharing research data</a> " (pp. 6-16) Leek, <a href="#">The Elements of Data Analytic Style</a> , chapters: "Tidying the data" and "Checking the data" (pp. 10-22) University of Leicester, <a href="#">Research Data Definitions</a> NCSU Libraries, <a href="#">Defining Research Data</a>	Lab 1
Week 3: Spatial, quantitative, qualitative, and "big" data   September 12th	
Force11, " <a href="#">Guiding Principles for Findable, Accessible, Interoperable and Re-usable Data</a> "	Lab 2

<p>Sutton et al., "<a href="#">A Gentle Introduction to GIS</a>"          USC Libraries, "<a href="#">Quantitative Methods</a>"          Sage, "<a href="#">Qualitative Research: Defining and Designing</a>" (pp. 1-17)          Ellingwood, Justin, "<a href="#">An Introduction to Big Data Concepts and Terminology</a>"</p>	
Week 4: Data Management   September 19th	
<p>Whyte, Angus and Jonathan Tedds, "<a href="#">Making the case for research data management</a>"          Akers, Katherine, and Jennifer Doty. "<a href="#">Disciplinary differences in faculty research data management practices and perspectives.</a>"          Wiener-Bronner, Danielle "<a href="#">Most Scientific Research Data From the 1990s Is Lost Forever</a>"          Perrier et al, "<a href="#">Research data management in academic institutions: A scoping review</a>"</p>	Lab 3
Week 5: Data cleaning and analysis   September 26th	
<p>Leek, <a href="#">The Elements of Data Analytic Style</a>, chapter: "Statistical modeling and inference" (pp. 34-44)          Nguyen, "<a href="#">Using Google Refine to clean messy data</a>"          Maceli, "<a href="#">Introduction to Text Mining with R for Information Professionals</a>"          Logan et al., "<a href="#">Choosing Statistical Software</a>"          Optional: Timmer, "<a href="#">Changing software, hardware a nightmare for tracking scientific data</a>"          BONUS: <a href="#">Programming Historian Lessons</a></p>	Lab 4
Week 6: Reproducibility   October 3rd	
<p>Steeves, Vicky, "<a href="#">Reproducibility Librarianship</a>"          Sayre, Franklin; Riegelman, Amy, "<a href="#">The Reproducibility Crisis and Academic Libraries</a>"          Vitale, Cynthia R.H. "<a href="#">Is Research Reproducibility the New Data Management for Libraries?</a>"          Dekker &amp; Lackie, "<a href="#">Technical Data Skills for Reproducible Research</a>" (pp. 93-112)</p>	Lab 5
Week 7: Legal and regulatory environment   October 10th	
<p>Boyle &amp; Jenkins, "<a href="#">The genius of intellectual property and the need for the public domain</a>" (pp. 10-14)          Arzberger et al., "<a href="#">An International framework to promote access to data</a>"          Hagedorn et al, "<a href="#">Creative Commons licenses and the non-commercial condition</a>"          Stodden, "<a href="#">The legal framework for reproducible scientific research: Licensing and copyright</a>"</p>	Work on check-in 1
Unit 2: Library Services	

Week 8: Data services in libraries   October 17th	
Goben, Zilinski, and Briney. " <a href="#">Going Beyond the Data Management Plan: Services and Partnerships.</a> " Salo, " <a href="#">Retooling libraries for the data challenge</a> " Reznik-Zellen et al., " <a href="#">Tiers of research data support services</a> " Emmelhain, " <a href="#">Data librarians in public libraries</a> " Coates, " <a href="#">Building data services from the ground up</a> "	Lab 6
Week 9: Data reference   October 24th	
Witt & Carlson, " <a href="#">Conducting a data interview</a> " Partlo, " <a href="#">The pedagogical data reference interview</a> " Carleton College, " <a href="#">Data, Datasets, and Statistical Resources</a> " Smith, Conte, and Guss, " <a href="#">Understanding Academic Patrons' Data Needs through Virtual Reference Transcripts</a> "	Project Check-in 1
Week 10: Data literacy & instruction   November 1st	
Shields, " <a href="#">Information literacy, statistical literacy, data literacy</a> " Rosenblum et al., " <a href="#">Collaboration &amp; co-teaching: Librarians teaching Digital Humanities in the classroom</a> " Kellam & Peter, " <a href="#">Data instruction: Statistical and data literacy</a> " Shorish, Yasmeen. " <a href="#">Data Information Literacy and Undergraduates: A Critical Competency</a> " Clement, Ryan, Blau, Amy, Abbaspour, Parvaneh, and Gandour-Rood, Eli, " <a href="#">Team-based data management instruction at small liberal arts colleges</a> "	
Week 11: Data collection services   November 7th	
Hogenboom et al., " <a href="#">Collecting small data</a> " Read through three of the <a href="#">Collections as Data Facets</a> projects	Work on your check-in!
Unit 3: Preservation, dissemination, and sustainability	
Week 12: Data sharing & publishing   November 15th	
NIH, " <a href="#">Frequently asked questions about the NIH Public Access Policy</a> " NSF, " <a href="#">Dissemination and sharing of research results</a> " Fienberg et al., <a href="#">Sharing Research Data</a> , " <a href="#">Issues and recommendations</a> " (pp. 3-32) Tenopir, C., Dalton, E.D., Allard, S., Frame, M., Pjesivac, I., Birch, B., Pollock, D. & Dorsett, K., " <a href="#">Changes in data sharing and data reuse practices and perceptions among scientists worldwide</a> " <a href="#">DataCite website</a> and specifically " <a href="#">Why cite data?</a> " ICPSR, " <a href="#">Phase 3: Best Practice in Creating Research Data</a> "; " <a href="#">Phase 6: Depositing Data</a> " (pp. 21-38)	2 <sup>nd</sup> final project check-in
Week 13: Data archives & repositories   November 21st	

<p>Kellam &amp; Peter, "<a href="#">Basic sources for supporting numeric data services</a>" (Read pp. 89-105; Skim interesting sources from pp. 106-149)</p> <p>Thiede, "<a href="#">Preservation in practice: A survey of New York City Digital Humanities practitioners</a>"</p> <p>Wilson, "<a href="#">How much is enough: metadata for preserving digital data</a>"</p> <p>Vines et al., "<a href="#">The availability of research data declines rapidly with article age</a>"</p>	
Week 14: NO CLASS   November 28th	
Enjoy your time off :)	
Week 15: Special concerns   December 6th	
<p>Cegłowski, "<a href="#">Deep-Fried Data</a>"</p> <p>Asher &amp; Jahnke, "<a href="#">Curating the ethnographic moment</a>" [PDF]</p> <p>Hurley, "<a href="#">When Academic Neurologists Leave, Who Owns Their Research?</a>"</p> <p>Moody, "<a href="#">Elsevier Says Downloading And Content-Mining Licensed Copies Of Research Papers 'Could Be Considered' Stealing</a>"</p> <p>Shaw &amp; Cloud, "<a href="#">Anonymization and microdata: Can we open up granular info without invading privacy?</a>"</p>	Work on final project in-class during lab time.
Week 16: Future/sustainability   December 13th	
<p>Timmer, "<a href="#">How science funding is putting scientific data at risk</a>"</p> <p>Goldstein &amp; Ratliff, "<a href="#">DataSpace: a funding and operational model</a>"</p>	Final projects presentations & hand in all final project materials.

### Textbooks, Readings and Materials

There is no required textbook. All readings and materials will be open access, so you will be able to read them without logging into anything.

### Assessment and Grading

#### Assignments & Projects

Labs (5% each)	30%
Participation	20%
Final project check-ins (10% each)	20%
Final project	30%
<b>TOTAL</b>	<b>100%</b>

#### Discussions & Participation (20%)

Each week, students should be prepared to discuss and/or ask and answer questions based on the readings or in-class exploratory labs. A student's participation grade will be based on facilitating class discussion during their assigned week (including coming up with 4-6

discussion questions) and actively participating in discussions led by other students. Discussion questions should be sent to the instructor before 11:45pm the Tuesday night before their assigned class. In addition to these in-class participation activities, some classes may require you to bring in outside materials/objects to fuel discussion.

### **Labs (6 @ 5% = 30%)**

The second half of the first 10 class periods will be devoted to hands-on activities designed to help you engage with that weeks' materials. Whatever is not finished in class will be assigned as homework and must be handed in by 11:55pm EST the Wednesday before the next class. These in-class labs are designed to underscore and amplify understanding in the lecture and readings for a given week. We will use software on the computers of the classroom, however all materials used (software, data, etc.) will be given at least a day in advance for any students who want to use their own machines.

### **Final Project (30%)**

You will have all semester to work on and refine a final project, which will be presented in the final class of the semester. You have two final projects to choose from, or you can present me an original idea for a final project by 5pm on Friday, September 7th. The format and length of presentations will be determined by the size of the class and the ratio of Project 1 choices to Project 2 choices. We will also have two graded project check-ins, to ensure all projects are on track.

*Project Choice 1 – Researcher Perspectives:* Design and carry out a small research project of your choice (focus on a data-informed study using either quantitative, qualitative, GIS data, or mixed methods). The end-product will be a research poster (online only, no need to print it) designed for a target conference, such as ACRL, SLA, RDAP or discipline-appropriate conference. Alongside the poster, you will also need to submit:

*Data management plan (2 pages max)*

Raw, analyzed, and publication-ready data

Data documentation or codebook (e.g. README and codebook if you're handing in a spreadsheet)

Any specialized analysis tools

1,000 - 1,500 word write-up on your study (with a bibliography or lit review section)

*Get some inspiration for posters here:*

<http://blogs.lse.ac.uk/impactofsocialsciences/2018/05/11/how-to-design-an-award-winning-conference-poster/>

<https://guides.nyu.edu/posters>

<https://researchguides.library.tufts.edu/c.php?g=344931&p=4823350>

*Project Choice 2 – Creating Data Services: based loosely off Dorothea Salo's Tool/Service Review project: <http://files.dsalo.info/668syll2014.pdf>*

First, you need to benchmark other services so you know where you want to build! Pick three institutions that currently offers data services, including data reference, data collection, research data services, and instruction services around data.

Services intended purpose and audience (e.g. patrons)  
Services fitness for purpose and audience  
Features (what problems does it solve? What gaps does it fill?)  
Limitations (what does it not do?)  
Prerequisites (what do users need to know/do before using the services?)  
Ease of use  
Future prospects (how trustworthy is the service?)  
Cost (staff, software, etc.)

This can be submitted as a spreadsheet. After you finish your peer benchmarking, you are expected to write a strategic plan on how to build and maintain your prospective data services department, including:

Outline of your organization's mission  
SWOT: your organization's strengths and weaknesses, as well as opportunities and threats  
4-5 goals aligned with mission  
Priorities, activities, objectives, strategies more in-depth

- Each goal should have a few different objectives/strategies associated with it

Roadmap and timeline  
1 page executive summary (should be written last!) to succinctly convey the future direction, priorities, and impact.

For inspiration, you can look at the [Strategic Agenda for Research Data Services](#) from Oregon State University.

### **Final Project Check-ins (20%)**

The first check-in will be on October 18th, and the second check-in will be on November 15th. These check-ins are to ensure that you are progressing on-schedule for your final projects, and to provide a way to get feedback at various points in the process from both the instructor and classmates.

#### *Project Choice 1 – Researcher Perspectives:*

1st Check-in – You will submit a 2 page data management plan and a 2 page (max) methodology statement to the instructor alongside any data that you've gathered or created. In-class, you are expected to present your thesis/idea, any data you have or plan on using, your methods, and your data management strategy. Your classmates will provide you feedback via a form created and circulated by the instructor.

2nd Check-in – By this point, you should be about 2/3 of the way done with your final project. As such, you'll hand in the data management plan (denoting any revisions from the DMP handed in for the first check-in and this one) and a 2 page (max) analysis statement that goes over how you've approached analyzing the data you've gathered or created. In-class, you'll present how far your project has come since the first check-in. Your classmates will provide you feedback via a form created and circulated by the instructor.

#### *Project Choice 2 – Creating Data Services:*

1st Check-in – You will submit your benchmarking study to the instructor ahead of class. In-class, you will present the results, including a discussion of how you chose to evaluate (e.g. on a scale of 1-4, completely qualitatively, etc.). Your classmates will provide you feedback via a form created and circulated by the instructor.

2nd Check-in – You will submit the outline of your organization’s mission, the 4-5 goals aligned with that mission, and the SWOT analysis to the instructor ahead of class. In-class, you will present this as if to a steering committee of a library, where you’d want to make the case for administrative buy-in for your goals and strategic mission. Your classmates will provide you feedback via a form created and circulated by the instructor.

*Project Choice 3 - Student-designed project:*

If you’ve chosen to design your own project, please discuss options for the check-in with the instructor.

**Pratt’s grading scale**

Superior work:	A 4.0 (96-100) A- 3.7 (90-95)
Very good work:	B+ 3.3 (87-89) B 3.0 (83-86) B-2.7 (80-82)
Marginally satisfactory:	C+ 2.3 (77-79) C 2.0 (73-76)
Failed:	F 0.0 (0-72)

**Portfolio**

Work completed for this course may be included in your portfolio. For more information on each program’s portfolio requirements, please visit the program’s respective webpage:

MS Library & Information Science: Portfolio - <http://bit.ly/prattmslisportfolio>

MS Information Experience Design: Portfolio - <http://bit.ly/prattmsixdportfolio>

MS Data Analytics and Visualization: Portfolio - <http://bit.ly/prattmsdavportfolio>

MS Museums and Digital Culture: Portfolio - <http://bit.ly/prattmsmdcportfolio>

You are encouraged to meet with your adviser about including projects in your portfolio.

**Pratt Institute-Wide Policies**

*This Course’s Attendance Policy*

When you have to miss a class, please notify the instructor in advance. Also, submit a 1-page reflection on the required readings for the missed class and complete the weekly lab that was missed. Students with 3 or more absences will be asked to drop the course, per Pratt policy. For more information on Pratt Institute’s Attendance Policy, please visit <http://bit.ly/prattattendance>.

*Academic Integrity Code*

Academic integrity at Pratt means using your own and original ideas in creating academic work. It also means that if you use the ideas or influence of others in your work, you must



acknowledge them. For more information on Pratt's Academic Integrity Standards, please visit <http://bit.ly/prattacademicintegrity>.

*Students with Disabilities*

Pratt Institute is committed to the full inclusion of all students. If you are a student with a disability and require accommodations, please contact the Learning/Access Center (L/AC) at [LAC@pratt.edu](mailto:LAC@pratt.edu) to schedule an appointment to discuss these accommodations. Students with disabilities who have already registered with the L/AC are encouraged to speak to the professor about accommodations they may need to produce an accessible learning environment.