



# VCU

Virginia Commonwealth University  
**VCU Scholars Compass**

---

VCU Libraries Faculty and Staff Publications

VCU Libraries

---

2016

## Data Management Roles for Librarians

Margaret E. Henderson

*Virginia Commonwealth University*, [mehenderson@vcu.edu](mailto:mehenderson@vcu.edu)

Follow this and additional works at: [http://scholarscompass.vcu.edu/libraries\\_pubs](http://scholarscompass.vcu.edu/libraries_pubs)



Part of the [Scholarly Communication Commons](#)

Publisher permission to use chapter.

---

Downloaded from

[http://scholarscompass.vcu.edu/libraries\\_pubs/40](http://scholarscompass.vcu.edu/libraries_pubs/40)

This Book Chapter is brought to you for free and open access by the VCU Libraries at VCU Scholars Compass. It has been accepted for inclusion in VCU Libraries Faculty and Staff Publications by an authorized administrator of VCU Scholars Compass. For more information, please contact [libcompass@vcu.edu](mailto:libcompass@vcu.edu).

## Chapter 14

### Data Management Roles for Librarians

#### In this Chapter

- Looking at data through different lenses
- Exploring the range of data use and data support
- Using data as the basis for informed decision making
- Treating data as a legitimate scholarly research product

Data is everywhere. Lots and lots of it. CERN has been releasing datasets from large hadron collider experiments. Data was first released in 2014 and some of it was used in a special high school physics class (Rao, 2014). The second release of data contained 300 TB of data (Liptak, 2016). Data and some tools can be downloaded by anyone from CERN Open Data (<http://opendata.cern.ch/>). By using predictive analysis on data from its 200 million customer accounts, a total of 1,000,000,000 GB of data on more than 1,400,000 servers, Amazon can provide personalized recommendations, and targeted marketing (Ann, 2016). HealthMap (<http://www.healthmap.org/>) processes data from tens of thousands of sources hourly to provide real-time information on disease outbreaks and their locations around the world. The information is free to everyone, and there is a mobile version of the site as well. eBird (<http://ebird.org/content/ebird/>) is a citizen science program that collects real-time data on bird sightings through an online checklist. The program has been running since 2002 and over 150,000 participants have submitted over 140,000,000 bird observations globally (Lagoze, 2014). Mark2Cure (<https://mark2cure.org/>) is a new citizen science project that identifies

concepts and concept relationships in biomedical text and analyzes them to find correlations and mentions, specifically for rare diseases, not easily found with regular indexing methods. This project relies on citizen science to conduct natural language processing tasks (Tsueng, 2016), From the largest research projects and companies to the rarest diseases that touch just a few people, data is important, and librarians can help manage that data. To paraphrase Shakespeare, “O brave new world, That has such data in ’t!” (The Tempest, Act 5, Scene 1).

## Data Librarianship in Context

### Data and Public Libraries

Public libraries are known for promoting information and digital literacy skills and providing lifelong learning opportunities (New York Comprehensive Center, 2012). Public libraries are also equalizing institutions, the place where people can go to get the information they need to be engaged in civic life, and the information is free, whether it be newspapers or magazines or databases provided by the library (Palfrey, 2015). Because of this, public libraries can play an important role in introducing community members to data and other technology issues that could impact their lives, such as open government data to petition for increased services or online privacy concerns. Public libraries are already helping with data literacy by bringing people together in many cities for hackathons (Alvarez, 2015) but also providing a venue for Citizen Science Day. The Citizen Science Day Expo at the La Jolla/Riford Library, part of the San Diego Public Library, included over 15 groups that work to encourage people to help with data collection or analysis (Tsueng, 2016). The La Jolla/Riford Library also has a

biology lab inside the library to allow people to do their own experiments and collect their own data (<http://lajollalibrary.org/your-library/bio-lab/>).

### Data and Business

Data is integral to business. High performing companies are more likely to use robust data mining and analytics on more data sources, which can be used to deliver better customer engagement and find efficiencies in business practices (Afshar, 2015). Decision making also relies on data. In the music industry, data can be used to inform decisions on when and where to release an album, or whether it should be streamed instead (Haugen, 2015). Business reliance on data impacts librarians who work in business and those who work helping business school students. Librarians with more knowledge of data sources and analytics will be better able to secure jobs in businesses, in knowledge management or the library, and academic business librarians need to be able to teach data information literacy skills to students. Business data has a spatial component as well, so an understanding of Geospatial Information Systems (GIS) is also important (Brody, 1999).

### Data and Citizen Science

Citizen science projects allow anyone to help as a data collector for a large scale project, such as the eBird project (<http://ebird.org/content/ebird/>) from Audubon and the Cornell Lab of Ornithology, or as a data analyst, like the people who help classify galaxies for Galaxy Zoo (<https://www.galaxyzoo.org/>). Zooniverse (<https://www.zooniverse.org/>) has gone beyond science to include many projects that include transcription of old letters, diaries, records, etc., so citizen science projects are becoming crowdsourced projects. This can have a direct impact on

libraries that are trying to digitize collections. The Library of Virginia's Transcribe project (<http://www.virginiamemory.com/transcribe/>) has multiple digital collections that are being transcribed by the public. The project is open access and the data will be machine readable, making the letters and other documents a perfect data source for future research.

### Data and Patient Care

As mentioned in Chapter 13, research reusing biomedical data is very important to the understanding of disease and advances in precision treatments. But there are other indirect ways data can be used. A data specialist was part of the Trafford Innovation and Intelligence Lab (<http://www.infotrafford.org.uk/>) group that used open data, mortality rates, obesity levels, rates of cardiovascular disease, and levels of physical activity in the area, to decide where to place defibrillators around the city. The same group used open and private data from area physicians to try and increase cervical cancer screening, and the result has been a 10 percent increase in screening (Ross, 2015). At a more personal level, precision medicine uses personal genomics data along with amassed data about genomics and treatments to find the best cancer treatments for individuals. Biomedical librarians work with health care practitioners to search for literature and genomic data, and the National Center for Biotechnology Information (NCBI) at the National Library of Medicine (NLM) is responsible for some of the databases that hold the genomic data and tools for sequence analysis. Consumer health librarians help patients find the information they need to understand the data and information coming from their physicians.

### Data and Health Care Information

Medical librarians are already involved in many teaching and research activities in hospital and health care education settings, including consumer health education; the privacy issues of patient health records make it a little harder to get involved with patient data, but it is possible. Many hospital systems and academic health care centers are pulling all patient data into an Enterprise Data Warehouse (EDW) to use for integration on multiple sites, management and strategic decision making, and clinical decision support (Evans et al., 2012). Medical and data librarians have the potential to use their reference and search skills, plus subject and database knowledge, to help health care practitioners search these databases for patient care and research information. The National Library of Medicine (NLM) has been heavily involved in working with genetic databases and tools to study sequences through the National Center for Biotechnology (NCBI), so many librarians started working with data and learning about database construction as a consequence. New York University Health Sciences Library (NYU HSL) worked with the NYU Department of Population Health to create a Data Catalog that would help researchers locate datasets relevant to their work, including information on licensing and faculty with expertise using the dataset (Lamb and Larson, 2016). The Data Catalog also contains information about datasets NYU researchers have collected, so their data can be shared. NYU HSL is making the code available in GitHub in the hopes that other libraries will set up data catalogs and information can be shared.

## Data and Open Science

As mentioned in the previous chapter, making data open, or at least publicly available, should help make science more reproducible. But there are many advocates who think the whole process of science should be done in the open to increase insights and transparency. Two recent

papers supporting open research and open access have been posting drafts on two different web sites and inviting comments as they edit and rework the paper, to show the value of doing work in the open (McKiernan, 2016; Tennant, 2016). The 2015 report, *Open Data in a Big Data World* (Science International), suggests that open science can help the scientific community have stronger dialogue and better engagement with wider society. The same report suggests that it is the ‘responsibility’ of libraries, archives, and repositories to develop and provide standards for data to ensure data is available and accessible over the long term for those who wish to use it.

### Data and Digital Humanities

Data is important to all types of research and scholarly projects, including areas outside of the sciences. Access to more data through computerized databases, 3-D scanning, and new imaging technologies are transforming research in archaeology. The Pottery Informatics Query Database (PIQD) archives 2D/3D-scanned ceramics, and it creates mathematical representations of shapes, allowing comparison between times and places, and allowing researchers to create a comprehensive ceramic typology for a region (Smith, 2012). And archaeologists are doing something libraries have been doing for a while – they are linking their databases of artifacts and other information into larger and larger networks so researchers can search for connections between places and times. The Digital Archeological Atlas of the Holy Land (<https://daahl.ucsd.edu/DAAHL/Home.php>) is part of Mediterranean Archaeological Network (MedArchNet; <http://medarchnet.org/>). People can manipulate maps, organize by time period, search by site or conditions, and review case studies. As with other disciplinary data management, librarians can help with ontologies, subject headings, database construction, and data curation.

While databases of archeological finds are similar to data collection in the sciences, digital humanities tries to fit data and information together in new ways, to help find new insights. “Hidden Patterns of the Civil War” (<http://dsl.richmond.edu/civilwar/>) is produced by the Digital Scholarship Lab in the Boatwright Memorial Library at the University of Richmond. The Yale University Library Digital Humanities Lab (<http://web.library.yale.edu/dhlab>) works on many projects, including Photogrammer (<http://photogrammar.yale.edu/>), which maps and organizes 170,000 photographs created by the United States Farm Security Administration and Office of War Information (FSA-OWI). Not only are library materials used in these projects, especially special collections materials, but library staff help with the projects as well, and librarians are making sure these projects are cataloged and findable.

#### Data and Social Good

The Opportunity Project (<http://opportunity.census.gov/>) was launched by the U.S. government in 2016 to encourage civic leaders, community organizations, and individual citizens to use open government data to increase opportunities for economic and social improvement. The website includes federal and local data, as well as tools to work with the data. Tools include Opportunity Score, that maps neighborhoods by proximity to jobs, and Transit Analyst, that maps transit routes in relation to community assets such as playgrounds, daycare, and health care. What Counts for America (<http://www.whatcountsforamerica.org/>) provides further information and case studies on open data and community improvements. These open resources and tools cover all subject areas, making it important for any librarian to be aware of disciplinary data resources.



Many non-profit organizations also need data librarians to manage the data gathered by funded research or collect open data to support the organization's mission. The International Food Policy Research Institute (IFPRI) Library (<http://library.ifpri.info/>) provides the usual library services of access to books, journals, databases, and datasets, but also helps IFPRI researchers with organizing their collected data and sharing that data with other repositories and academic researchers. They also offer training to researchers on data visualizations, tools, and widgets.

#### Data and Education

Parents and students need to be aware of extensive data collection at all levels of education. Students are tracked throughout a student's time in a K-12 system, and these systems are now sharing information. Some testing and data collection can only be done with parental permission, but other information is automatically collected if a child is in public schools (Strauss, 2015). Higher education also has many systems compiling information and keeping records about students. Learning analytics is one of the systems used in many places to help students with learning objectives and help teachers create customized education plans. Privacy concerns have slowed the uptake of some learning analytics, and policies are being developed to ensure that privacy, anonymity, and data security concerns are addressed by the systems any school uses (Drachler and Greller, 2016). Librarians, especially those with teaching roles, need to be aware of these issues when using online learning management systems. And data librarians that support educational research need to carefully screen any databases or datasets they are asked to support.

## Data and Privacy

Internet privacy is a big concern for many people. Businesses collect data on what people buy or watch, and many people appreciate the recommendations or deals they get later based on that data. But there are things that people wish to keep private, and at times it can be hard to remain anonymous on the Internet. When does the collection of data about an individual by government or companies become a problem? Even blogs, which on the surface seem to be a public data source, have private aspects that must be considered before mining them for research (Eastham, 2011). Anonymization or deidentification needs to be carefully planned, but at the same time, studies that profess to re-identify people from a given data set need to be scrutinized and proper studies conducted to make sure any problems are resolved (Barth-Jones, 2014). The American Library Association (ALA) has Privacy as a Core Value (<http://www.ala.org/advocacy/intfreedom/statementspols/corevalues>). As much as librarians want to support open and linked data, it is important to remember the obligation to protect users' privacy and only provide data to others that is consistent with this value (Campbell and Cowan, 2016).

## Data and Assessment

Libraries regularly collect data to assess services and user needs. Reference service data has been used to redistribute workload and provide data analysis experience for librarians (Goben and Raszewski, 2015). A database of librarian-mediated literature searches was analyzed by subject and purpose to help with librarian training and to allow for proactive preparation of materials, as well as workload tracking (Lyon, et al., 2014). Collecting and analyzing surveys to discover the data needs of researchers is an important step in setting up research data

management services (Norton et al., 2016; Whitmire, Boock, and Sutton, 2015; Weller and Monroe-Gulick, 2015).

Librarians also possess the skills to be involved in bibliometric assessment to help researchers understand the various metrics that measure the impact of all types of scholarly research output. While bibliometrics have usually been based on data collected about citation patterns, it is important for librarians to also learn about the theory behind these metrics if they are going to provide impact metrics as a support service (Corrall, Kennan, and Afzal, 2013). Understanding the new citation-independent metrics that are based on usage, both downloads and page visits, and mentions of research in blogs, newspapers, and other venues is also important. Non-traditional scholarly outputs, such as data, software, and blogs, are starting to be considered as part of grant applications and promotion and tenure reviews, and data librarians need to be aware of metrics that can help measure the impact of these types of scholarly products (Khodiyar, Rowlett, and Lawrence, 2014). Thanks to grant funding, Impact Story (<https://impactstory.org/>) is providing a free service to track all these metrics for researchers with an ORCID (Open Researcher and Contributor ID, <http://orcid.org>). The skills of a librarian – knowledge of authority control for name disambiguation and subject ontologies – are ideal for helping with academic professional metrics (Lowe, 2013). This is especially true for researchers in arts and humanities, where there is not the same history of metric usage (Priego, 2015).

### Data and Scholarly Communication

Data is neither fish nor fowl when discussions of scholarly communications come up. You can't copyright data, but you can license it. Data is not usually owned by the person who collects it and manages it; it is owned by the employer/funder. Data is not really considered a full

scholarly product by most people, like an article or book is, and yet, “Data are outputs of research, inputs of scholarly publications, and inputs to subsequent research and learning. Thus they are the foundation of scholarship” (Borgman, 2007: 115). As funding agencies start to recognize data as a research product, and publishers start to require data to be linked to papers, making sure data is cited properly will help increase the visibility of data and encourage more people to share their data. Data librarians, especially those involved with data curation, can help by making sure data has a DOI (digital object identifier, <https://www.doi.org/>) whenever possible and helping with the organization of large datasets so specific parts of the dataset can be cited separately (Callaghan, 2014). Data librarians should be aware not only of funder and publisher policies, but also policies and recommendations put out by disciplinary organizations (Murphy, 2014), and interdisciplinary groups that represent the data and publishing interests of researchers, such as the Research Data Alliance (<https://rd-alliance.org/>) or Force11 (<https://www.force11.org/>).

### Data and Public Access Policies

Keeping up with all the various institutional and funder policies and rules, may seem to take up a lot of the time of a data librarian, even with the new SPARC site collecting all the OSTP policies together (<http://datasharing.sparcopen.org/>). However, it is important to remember that helping researchers comply with applicable policies requiring data management plans and public access to data is not an end in itself. The policies are there to encourage a new open way to conduct research that will benefit everyone. Helping with compliance will only lead to a sustainable data service if data librarians educate as they help. Most publishers will streamline public access for papers and digital data, just as they have for the NIH Public Access

Policy. But now, while there is an opportunity, data librarians need to stress the copyrights researchers are giving away and educate them on data ownership so they don't go against institutional policies. It won't be long before authors/researchers go through the process of publication without a thought to copyright, just as they do now, because there will be boxes to check for funders and an upload button for data deposit, so use this transition period to educate. This is especially true with documentation of deposited data. PubPeer (<https://pubpeer.com/>) has shown that there are many people out there studying papers, looking for discrepancies. If documentation doesn't clearly explain all the data, a retraction could ensue. This is where data librarians can find sustainable work – setting up work flows and data entry forms, helping with data dictionaries, helping capture code for analysis, making sure spreadsheets have a data dictionary to explain every row and column. Facing a retraction is a huge stick, and this kind of data management work is not spelled out in policies.

### Finding a Role for Librarians

The examples and suggestions above give some ideas about how data librarians can work with a wide range of researchers in all types of libraries. Data has become a part of every discipline and everyday living, whether people realize it or not. Some of the people who have to deal with data already know how to collect, organize, describe, store, analyze, visualize, and share their data, but most people don't know how to do it all and could use some help. And there are also people who have no idea where to start. Data is information, something librarians have always worked with, and the skills needed to help people with data are the same as those used to help people find a book, website, government agency address, article, video, or children's book they need. But while children learn to read in school, very few students learn about statistics, let

alone data information literacy in high school or college. Data librarians can guide people to the data they need, but they also need to be ready to help people understand the information they have received because it will be in a form most people don't recognize.

Data librarianship is not particularly new. In 2004, Ann S. Gray, Data Reference Librarian at Princeton University Library, was recommending that librarians learn more about data and statistics so they could extend their role beyond acting as intermediaries to statistical resources. But, despite the recognition that data is a hot topic for librarianship (Laguardia, 2015), and some large research university library directors planning to integrate data management into most university librarians' roles (Marcum, Schonfeld, Thomas, 2015), the majority of ACRL institutions surveyed are not offering research data services (Tenopir, et al., 2015). Interviews with ACRL directors (Tenopir et al., 2015), indicated that many were concerned about the technical capacity of libraries and need for technology training for staff, but it might be more helpful to consider data management as a research service, not a technical service. Sayeed Choudhury, Associate Dean for Research Data Management, Johns Hopkins University Sheridan Libraries, encouraged librarians to promote themselves as providing research support, not just a library service (Choudhury, 2013). Data management should also be considered in the larger context of the institution. A survey of health sciences library directors found that libraries need to align with the institutional mission, and librarians need to "become more involved in the fabric of the institution so that they can anticipate where unique skill sets might be required to foster an evidence-rich environment" (McGowan, 2012: 44).

The National Library of Medicine has been awarding NLM Administrative Supplements for Informationist Services in NIH-funded Research Projects (<https://www.nlm.nih.gov/ep/AdminSupp.html>) since 2012 to explore the possibility of

librarians, especially those that specialize in data, becoming integral to the research process. Several of the librarians from these projects have written about their experiences working with a research team, which include creating a data dictionary that allowed all team members to discuss and request data efficiently (Gore, 2013), recommending tools and workflows to help with the collection of data and specimens (Surkis et al., 2013), and helping with data management and curation (Federer, 2013; Hanson et al., 2013). These librarians show that there is a place for librarians in research, and the focus does not always need to be highly technical. In the initial assessment of the informationist program at the NIH Library (Robison, Ryan, and Cooper, 2009), informationists not saved time by providing expert information retrieval and training, but they also acted as a portal to collaborators, partners, and contacts. This idea of librarian as facilitator is important when it comes to connecting researchers with the data infrastructure they need within or outside of the institution.

Elaine Martin (2015), Director of Library Services University of Massachusetts Medical School, is concerned about the hesitancy on the part of librarians and library administration to participate in the data movement. She has proposed a framework for the librarian's role in data management that centers around more than just the technological aspects of data management (see Figure 14.1). Librarians already help out in the areas of services, best practices, literacy, archives and preservation, and policy. Martin also points out that the values library science brings to data – a focus on the user, user needs, and user behavior; an ethical base; collaboration; and equal access to information – are unique among the disciplines that are involved in data science, such as computer science or informatics.

<insert Figure 14.1 here>

Figure 14.1. User-Centered Data Management Framework for Librarians. Building a community of data science librarians to evolve the data science discipline. (Martin, 2015)

Used with permission by Elaine Martin

Providing data management help to library users does not need to involve expensive infrastructure. It does require an entrepreneurial outlook – somebody who is willing to sell the service and adapt to community needs if feedback suggest a better way to provide support. There will always be researchers and librarians who do not see a role for the library in data management. However, there will be many more researchers who are happy to have help and guidance on data management, organization, and policies. And there are many librarians and subject specialists working in libraries who see the need for data management support for the researchers and faculty they interact with at their institutions. Data is a big tent and there is room for many different specialties and disciplines, and there is a need for them as well. With careful planning and preparation, including continuing education, librarians can definitely be involved in data management.

Data management is a moving target right now and will be for many years to come as all the stakeholders, researchers, governments, funders, academic institution, companies, etc., hash out standardized ways to work with data. There may never be one right way to process a dataset, but there are efforts being made to at least bring certain types of data together in a standard way, for example health records, climate data, or genetic data. From these efforts, it is easy to see various ways for groups to work together to make data usable. In the meantime, librarians and other data professionals will need to include time for continuing education in their schedules to



keep up with best practices, new policies and regulations, changes in scholarly communication, and open access and licensing practices.

### Key Points

Technology has made it easy to collect and acquire data about almost anything, but it requires training or help to make sense of it all.

- Collecting and using data is integral to many disciplines.
- Libraries and librarians have the skills needed to help with data management.
- Data can be used to help with social and community concerns, as well as research.
- The field of data management is constantly changing so lifelong learning is essential for anyone helping with data.

Hopefully this book will inspire librarians to start providing data management services at their institutions.

### References

Afshar, Vala. 2015. "2015 State of Analytics - 20 Key Business Findings." *The Blog, the Huffington Post*. [http://www.huffingtonpost.com/vala-afshar/2015-state-of-analytics-2\\_b\\_8611382.html](http://www.huffingtonpost.com/vala-afshar/2015-state-of-analytics-2_b_8611382.html).

Alvarez, Barbara. 2015. "Hackathons @ the Library." *Public Libraries Online* [blog] <http://publiclibrariesonline.org/2015/05/hackathons-the-library/>.

Ann. 2016. *How Amazon Uses its Own Cloud to Process Vast, Multidimensional Datasets*.

DZone [website] <https://dzone.com/articles/big-data-analytics-delivering-business-value-at-am>.

- Barth-Jones, Daniel. 2014. "The Antidote for "Anecdotal": A Little Science can Separate Data Privacy Facts from Folklore." *Info/Law* [blog]  
<https://blogs.harvard.edu/infolaw/2014/11/21/the-antidote-for-anecdotal-a-little-science-can-separate-data-privacy-facts-from-folklore/>.
- Borgman, Christine L. 2007. *Scholarship in the Digital Age: Information, Infrastructure, and the Internet*. Cambridge, MA: MIT Press.
- Brody, Roberta. 1999. "Geographic Information Systems: Business Applications and Data." *Journal of Business & Finance Librarianship* 5, no. 1: 3-18. doi:10.1300/J109v05n01\_02
- Callaghan, Sarah. 2014. "Preserving the Integrity of the Scientific Record: Data Citation and Linking." *Learned Publishing* 27: 15-24. doi:10.1087/20140504
- Campbell, D. Grant, and Scott R. Cowan. 2016. "The Paradox of Privacy: Revisiting a Core Library Value in an Age of Big Data and Linked Data." *Library Trends* 64, no. 3: 492-511.
- Choudhury, Sayeed. 2013. "Open Access & Data Management are do-Able through Partnerships." ASERL (Association of Southeastern Research Libraries) Summertime Summit, Atlanta, GA, August 8. <https://smartech.gatech.edu/handle/1853/48696>.
- Corrall, Sheila, Mary Anne Kennan, and Waseem Afzal. 2013. "Bibliometrics and Research Data Management Services: Emerging Trends in Library Support for Research." *Library Trends* 61 (3): 636-674.
- Drachsler, Hendrik and Wolfgang Greller. 2016. "Privacy and Analytics: It's a DELICATE Issue a Checklist for Trusted Learning Analytics." Proceedings of the Sixth International Conference on Learning Analytics & Knowledge, LAK '16. Edinburgh, United Kingdom, ACM, doi:10.1145/2883851.2883893.

- Eastham, Linda A. 2011. "Research using Blogs for Data: Public Documents Or Private Musings?" *Research in Nursing & Health* 34, no. 4: 353-361. doi:10.1002/nur.20443
- Evans, R. S., J. F. Lloyd, and L. A. Pierce. 2012. "Clinical use of an Enterprise Data Warehouse." *AMIA Annual Symposium Proceedings* 2012: 189-198.
- Federer, Lisa. 2013. "The Librarian as Research Informationist: A Case Study." *Journal of the Medical Library Association : JMLA* 101, no. 4: 298-302. doi:10.3163/1536-5050.101.4.011
- Goben, Abigail and Rebecca Raszewski. 2015. "The Data Life Cycle Applied to our Own Data." *Journal of the Medical Library Association : JMLA* 103, no. 1: 40-44. doi:10.3163/1536-5050.103.1.008
- Gore, Sally A. 2013. "A Librarian by any Other Name: The Role of the Informationist on a Clinical Research Team." *Journal of eScience Librarianship* 2, no. 1: Article 6.  
<http://dx.doi.org/10.7191/jeslib.2013.104>.
- Gray, Ann S. 2004. "Data and Statistical Literacy for Librarian." *IASSIST Quarterly* 282, no. 3: 24-29. [http://www.iassistdata.org/sites/default/files/iq/iqvol282\\_3gray.pdf](http://www.iassistdata.org/sites/default/files/iq/iqvol282_3gray.pdf).
- Hanson, Karen L., Theodora A. Bakker, Mario A. Svirsky, Arlene C. Neuman, and Neil Rambo. 2013. "Informationist Role: Clinical Data Management in Auditory Research." *Journal of eScience Librarianship* 2, no. 1: Article 7. <http://dx.doi.org/10.7191/jeslib.2013.1030>
- Haugen, Brad. 2015. "Data Will Save Music." *Tech Crunch*.  
<http://techcrunch.com/2015/02/17/data-will-save-music/>.
- Khodiyar, Varsha K., Karen A. Rowlett, and Rebecca N. Lawrence. 2014. "Altmetrics as a Means of Assessing Scholarly Output." *Learned Publishing* 27: 25-32.  
doi:10.1087/20140505

- Lagoze, Carl. 2014. "eBird: Curating Citizen Science Data for use by Diverse Communities." *International Journal of Digital Curation* 9, no. 1: 71-82.
- LaGuardia, Cheryl. 2015. "Where are we Headed? an Unscientific Survey." *Library Journal* 140, no. 19 (November 15): 14.
- Lamb, Ian and Catherine Larson. 2016. "Shining a Light on Scientific Data: Building a Data Catalog to Foster Data Sharing and Reuse." *Code{4}Lib* 32 (April 25).  
<http://journal.code4lib.org/articles/11421>.
- Liptak, Anderw. 2016. "CERN just Dropped 300 Terabytes of Raw Collider Data to the Internet." *Gizmodo*. April 23. <http://gizmodo.com/cern-has-released-300-terabytes-of-collider-data-to-the-1772642139>.
- Lowe, David B. 2013. *A Visible Job to do: Some Thoughts on Opportunities for Libraries Concerning Academic Professional Metrics*: UConn Libraries Published Works. Paper 46.  
[http://digitalcommons.uconn.edu/libr\\_pubs/46](http://digitalcommons.uconn.edu/libr_pubs/46).
- Lyon, J. A., R. Garcia-Milian, H. F. Norton, and M. R. Tennant. 2014. "The use of Research Electronic Data Capture (REDCap) Software to Create a Database of Librarian-Mediated Literature Searches." *Medical Reference Services Quarterly* 33, no. 3: 241-252.  
doi:10.1080/02763869.2014.925379
- Marcum, Deanna, Roger Schonfeld, and Sarah Thomas. 2015. *Office of Scholarly Communication: Scope, Organizational Placement, and Planning in Ten Research Libraries*: Ithaka S+R; Harvard Library. <http://www.sr.ithaka.org/publications/office-of-scholarly-communication/>.
- Martin, Elaine R. 2015. "The Role of Librarians in Data Science: A Call to Action." *Journal of eScience Librarianship* 4 (2): e1092. <http://dx.doi.org/10.7191/jeslib.2015.1092>.

- McGowan, J. J. 2012. "Tomorrow's Academic Health Sciences Library Today." *Journal of the Medical Library Association : JMLA* 100, no. 1: 43-46. doi:10.3163/1536-5050.100.1.008
- McKiernan, Erin, Philip E. Bourne, C. Titus Brown, et al. 2016. "The Benefits of Open Research: How Sharing can Help Researchers Succeed." *Figshare* preprint. <https://dx.doi.org/10.6084/m9.figshare.1619902.v5>.
- Murphy, Fiona. 2014. "Data and Scholarly Publishing: The Transforming Landscape." *Learned Publishing* 27: 3-7. doi:10.1087/20140502
- New York Comprehensive Center. 2012. *Public Libraries Informational Brief: Impact of Public Libraries on Students and Lifelong Learners* . New York: New York Comprehensive Center Educational Technology Team. [http://www.nysl.nysed.gov/libdev/nyla/nycc\\_public\\_library\\_brief.pdf](http://www.nysl.nysed.gov/libdev/nyla/nycc_public_library_brief.pdf).
- Norton, Hannah F., Michele R. Tennant, Cecilia Botero, and Rolando Garcia-Milian. 2016. "Assessment of and Response to Data Needs of Clinical and Translational Science Researchers and Beyond." *Journal of eScience Librarianship* 5, no. 1: e1090. <http://dx.doi.org/10.7191/jeslib.2016.1090>.
- Palfrey, John. 2015. *BiblioTECH: Why Libraries Matter More than Ever in the Age of Google*. New York: Basic Books.
- Priego, Ernesto. 2015. #HEFCEMetrics: More on Metrics for the Arts and Humanities Dr Ernesto Priego [blog]. <https://epriego.wordpress.com/2015/01/16/hefcemetrics-more-on-metrics-for-the-arts-and-humanities/>.
- Rao, Achintya. "CMS Releases First Batch of High-Level LHC Open Data." *CERN* [website] last modified 2014-11-20, <http://cms.web.cern.ch/news/cms-releases-first-batch-high-level-lhc-open-data>.

- Robison, R. R., M. E. Ryan, and I. D. Cooper. 2009. "Inquiring Informationists: A Qualitative Exploration of our Role." *Evidence Based Library and Information Practice* 4, no. 1: 4-16.
- Ross, Eleanor. 2015. "How Open Data Can Help Save Lives." *The Guardian*, Tuesday 18 August. <http://www.theguardian.com/media-network/2015/aug/18/open-data-save-lives-emergency-services-disaster-relief>.
- Science International. 2015. *Open Data in a Big Data World*. Paris: International Council for Science (ICSU), International Social Science Council (ISSC), The World Academy of Sciences (TWAS), InterAcademy Partnership (IAP). <http://www.icsu.org/science-international/accord/open-data-in-a-big-data-world-long>.
- Smith, Neil G., Avshalom Karasik, Tejaswini Narayanan, Eric S. Olson, Uzy Smilansky, and Thomas E. Levy. 2014. "The Pottery Informatics Query Database: A New Method for Mathematic and Quantitative Analyses of Large Regional Ceramic Datasets." *Journal of Archaeological Method and Theory* 21, no. 1: 212-250. doi:10.1007/s10816-012-9148-1
- Strauss, Valerie. 2015. "The Astonishing Amount of Data being Collected about Your Children." *The Washington Post/Answer Sheet* November 12. <https://www.washingtonpost.com/news/answer-sheet/wp/2015/11/12/the-astonishing-amount-of-data-being-collected-about-your-children/>.
- Surkis, Alisa, Aileen McCrillis, Richard McGowan, et al. 2013. "Informationist Support for a Study of the Role of Proteases and Peptides in Cancer Pain." *Journal of eScience Librarianship* 2, no. 1: Article 9. <http://dx.doi.org/10.7191/jeslib.2013.1029>.
- Tennant, Jonathan P., Francois Waldner, Damien C. Jacques, Paola Masuzzo, Lauren B. Collister, and Chris H. J. Hartgerink. 2016. "The Academic, Economic and Societal Impacts

- of Open Access: An Evidence-Based Review” [Version 1; Referees: 2 Approved, 1 Approved with Reservations]. *F1000Research* 5: 632. doi: 10.12688/f1000research.8460.1
- Tenopir, Carol, Elizabeth D. Dalton, Suzie Allard, et al. 2015. "Changes in Data Sharing and Data Reuse Practices and Perceptions among Scientists Worldwide." *PLoS ONE* 10 (8): e0134826. <http://dx.doi.org/10.1371/journal.pone.0134826>.
- Tsueng, Ginger. 2016. "Citizen Science Day Expo." *The Su Lab* [Blog] March 11 (updated April 8). <http://sulab.org/2016/03/citizen-science-day-expo/>.
- Tsueng, Ginger, Max Nanis, Jennifer Fouquier, Benjamin Good, and Andrew Su. 2016. "Citizen Science for Mining the Biomedical Literature." *bioRxiv* 038083. doi:10.1101/038083
- Weller, Travis and Amalia Monroe-Gulick. 2015. "Differences in the Data Practices, Challenges, and Future Needs of Graduate Students and Faculty Members." *Journal of eScience Librarianship* 4, no. 1. <http://dx.doi.org/10.7191/jeslib.2015.1070>.
- Whitmire, Amanda Lea, Michael Boock, and Shan C. Sutton. 2015. "Variability in Academic Research Data Management Practices: Implications for Data Services Development from a Faculty Survey." *Program: Electronic Library and Information Systems* 49, no. 4: 382-407. <http://dx.doi.org/10.1108/PROG-02-2015-0017>.