

Personal Statement

Ryan P. Womack

August 8, 2018

Introduction

My approach to librarianship is to seek the areas where I can most effectively apply my most advanced skills through research, teaching, and service. Over the years, that has meant a shift in focus from business to data as I have developed statistical skills to complement my original background in economics. I have also sought to take on appropriate leadership roles for my areas of speciality, and to make creative contributions at many levels, from local to national and international. I have endeavoured to make the products of my work widely and openly accessible so that my writings, code, video, and other instructional material can be used by any audience. In my view, the fundamental ideals of librarianship are providing access to information and teaching users of information the tools that can convert information into knowledge, enabling users to embark on a journey of lifelong independent learning. In my work, the “information” is primarily numeric data, and the “tools” are data archives and software for statistical and numerical analysis. The context of data librarianship is different, but the ideals are the same as elsewhere in librarianship. I also seek to push boundaries by taking on new areas and approaches that have not often been seen in libraries before.

[Note: throughout this document, references to supplemental materials are indicated by *superscripts* as follows: letter-number combinations such as “A5” or “P7” refer to Scholarship. Tab number references such as “Tab 3” refer to Librarianship and Service supplemental materials.]

Librarianship

Business From 1999 to 2007, my position in Rutgers University Libraries (RUL) was *Business and Economics Librarian*. As such, I liaised with the Rutgers Business School on the New Brunswick campus, and worked closely and collaboratively with business librarians at Rutgers-Newark. A major focus during the early 2000s was combining resources to convert local print collections (which were being paid for separately by the business collections at Newark, New Brunswick, and Camden) into shared electronic databases. The print journals and massive runs of print handbooks containing corporate information gave way to online databases, leading to the systemwide purchase of such resources as Mergent, ReferenceUSA, Factiva, and Business Source Premier. During this period, I also participated in Collection Development Council and other responsibilities typical for subject liaisons.

I provided instruction and outreach for courses in business and economics, including sessions focused on specialized corporate and financial resources such as COMPUSTAT, along with regular

reference on business and general topics, transitioning business users to the rapidly developing capabilities of new electronic resources.

Throughout my career at RUL, I have made use of my academic training in economics to serve as subject specialist and liaison to the Department of Economics in the School of Arts and Sciences, and also to the Department of Agricultural, Food, and Resource Economics in the School of Environmental and Biological Sciences.

My favorite aspect of business librarianship had always been working with financial data and statistics. So, when the opportunity arose to move into a role focused on data, I took it, taking the title *Data and Economics Librarian* in 2008. I maintained my role as liaison to Economics and Agricultural Economics, while relinquishing my business responsibilities. As the sole librarian for data in New Brunswick, I assumed responsibility for all aspects of data-related public services.

Research Data Beginning in 2009, RUL began its response to the growing interest in managing research data with the establishment of the RUcore Data Working Group. I was a member and leading contributor to all of the Libraries' systemwide research data initiatives over the years from 2009 to 2016 as the working group grew into the RUresearch Data Team, which I chaired from 2011 to 2016. Along with colleagues in **RUcore** (the Rutgers University Community Repository) and across the Libraries, we developed a working data portal built from the advanced Fedora repository architecture that had been created for RUcore. We then began working with Rutgers faculty to archive and share their research datasets, overcoming many technical challenges along the way in order to develop advanced functionality for preserving and presenting data, and gaining hands-on knowledge of the research data management process. I helped train a large cohort of librarians as part of a special course in handling research data in 2012 and again in 2014, as the Data Team grew to include more than 25 librarians and staff.^{Tab 4}

During this initial growth phase, I served as Research Data Manager for RUcore in addition to my other duties. During 2011 to 2013, I led outreach on campus about the Libraries' data services, reaching over 100 faculty and generating an initial group of about 10 data deposits. The successful implementation of the data portal was instrumental in the appointment of a full-time Research Data Manager who I then trained.

In 2013, along with a change in title to *Data Librarian*, my duties were adjusted to reflect a greater emphasis on public services work—assisting users with finding and using data—and to using my expertise in statistical and numeric data to contribute to other projects. I continued to lead the RUresearch Data Team until it was eliminated in 2016. I currently contribute my research data expertise to the New Brunswick Libraries Research Data Outreach Team and the NSF-funded **Virtual Data Collaboratory** (VDC) (see <https://rdi2.rutgers.edu/page/virtual-data-collaboratory> and Tab 2). The NBL Research Data Outreach Team offers research data management consulting services to the Rutgers-New Brunswick community. The VDC is being built as a shared infrastructure for researchers to collaborate on big data among Rutgers, Penn State, Temple, and other regional institutions. As Senior Personnel on this project, I advise on tools for data analytics and educate users about research data management.^{Tab 3, Tab 5}

Data Public Services My public services work with data and statistical resources has become the core of my library work. As Data Librarian, I develop and support user-centric data services in response to the needs of the Rutgers community. I liaise and consult with faculty and students

on a wide range of data needs (finding data, selecting databases, working with datasets and software). In recent years, I have answered around 200 individual data-specific questions per year (a high of 250 in 2014-15, 176 to date in 2017-2018), questions that come to me because of my data expertise. These relationships, which often start with e-mail queries, can develop into long-term consulting on research topics. I give course-based instruction on subject-specific data resources in collaboration with other liaison librarians, addressing the data needs of many social science disciplines, such as Political Science, Social Work, and others.^{Tab 3, Tab 6}

I deliver a continually updated series of workshops every semester, which include introductions to statistical software (R, SAS, Stata, SPSS), and more advanced topical sessions on reproducible research, time series analysis, data visualization, multivariate statistics, and beyond. These workshops attract graduate students and faculty from across a broad range of disciplines. Each year, I deliver approximately 20 of these sessions. In 2017-2018, I delivered 18 workshops attracting 226 students. These sessions, particularly those involving the use of R software, have become the most requested and popular aspect of my work with the Rutgers community. In addition to open public workshops in the Libraries, I have been asked to repeat these workshops in the Graduate School of Education, the Bloustein School of Planning and Public Policy, the Professional Master's Program and several other venues. Data can be involved in almost any discipline, and it has been exciting and gratifying to be able to offer useful expertise to groups as disparate as the School of Management and Labor Relations and the Center of Alcohol Studies.^{Tab 7}

As the primary public services librarian for data in New Brunswick, with a population of over 50,000 students, meeting the needs of users is sometimes overwhelming. In response, I have been creative in developing services that can leverage my expertise in ways that serve a broad community. As further discussed under Scholarship, I have created online materials for most of my workshop series, including video screencasts and reusable code, so that they are available to anyone at any time.

I maintain research guides highlighting the most significant data resources (libguides.rutgers.edu/data and libguides.rutgers.edu/databysubject), as well as research data management and economics guides.^{Tab 8} I manage data-related collections and acquire datasets as needed, from widely used databases such as Data-Planet Statistical Datasets to unique data resources requested by Rutgers users such as the Vietnam Household Living Standards Survey. I am Rutgers University's official representative (OR) to the largest social science data archive in the world, the Inter-university Consortium for Political and Social Research (ICPSR) based at the University of Michigan. I am also the official representative to the Roper Center, the largest collection of public opinion data and polls. As the OR, I am the primary support person on campus for outreach and assistance in using these data archives that are essential to research in the social sciences and beyond.

I initiated the creation of the *Secure Data Facility* in 2011 at Alexander Library, providing a secure space and non-networked PC that meets the protocols required for access to restricted federal government data sources. This space has seen a steady flow of doctoral students using it since its creation, in some cases for research that is central to their dissertations. This service brought Rutgers up to the level of service of our peer research institutions. Successful operation of this facility led to the funding of an upgrade to the key system to a fully electronic card/PIN reader by the School of Arts and Sciences and the Department of Economics. I also consult with researchers about IRB, data security and data privacy issues.^{Tab 9}

In order to deepen my expertise, I completed a *Master of Science in Statistics and Biostatistics* at

Rutgers in 2015. This has enabled me to help users tackle thornier questions and more complicated data problems, while providing a more nuanced and scientific approach to analytical issues.^{Tab 10}

Leadership Within the New Brunswick Libraries (NBL), I have held leadership roles, such as a term as Chair of the New Brunswick Collections Group (2009-2011). I was selected for the 2015 class of the the Triangle Research Libraries Network Management Academy at UNC-Chapel Hill, and attended this competitive program for leadership development.^{Tab 11}

Most notably, I played a key role in the reorganization of NBL, leading teams during the planning and conceptualization of the organizational structure during 2015, and then becoming the first *Head of the Research and Content Department* (2016-present), supervising teams for Content, Graduate Services, Research Data, Research Spaces, and Scholarly Communication. This reorganization was intended to bring faculty librarians and staff together to address emerging needs within NBL. In this role, I seek to guide librarians and staff to develop effective services supporting Rutgers-New Brunswick researchers. Developing an efficient team structure, setting goals, working with team leaders, and supporting them through budget and personnel are all part of this work.^{Tab 12}

In order to address the need for consulting and training in advanced research methodologies that I have encountered in my data work, and others in the Libraries encounter in their fields, I conceptualized and led the pilot *Graduate Specialist Program* in 2018, which hired graduate student experts in data science and digital humanities to deliver workshops and provide one-on-one assistance in these methodologies. In the initial pilot, this enabled the Libraries to offer new workshops in Python for data science and text analysis tools for the digital humanities. This program will expand in Fall 2018 to include four specialists for quantitative data analytics, qualitative methods, digital humanities, and open science and reproducible research.^{Tab 13}

Outreach and Teaching A strong basis of expertise in statistics and statistical resources has enabled me to extend outreach on a number of levels. In order to meet the demands of a large and dispersed user base, I have from the inception of my work as Data Librarian used all of the technology at my disposal to make my instructional materials available on the web, so that it can be accessed at any time, anywhere, on demand.

My strategy and philosophy for developing services is to focus on areas where there is high demand that match my skillset. For example, since beginning my first introductory workshop on R statistical software, I have had repeated calls for more offerings like it. This has led to investigations of Data Visualization, Big Data, and several other topics, all of which resulted in workshop presentations for both Rutgers and external users.

Since there is a limit to the number of in person sessions I can provide, I began creating video versions of the workshops so that they would be available for reuse. These screencasts are hosted on my YouTube channel (youtube.com/librarianwomack), which has amassed over 2,100 global subscribers and 310,000 views at the time of writing. I maintain a blog with over 11,000 views (ryandata.wordpress.com) and a Twitter account (twitter.com/ryandata) in order to disseminate announcements and data-related discussion. I also distribute software code, presentation files, data, and documentation for my workshops through my Github account (github.com/ryandata), which provides a highly flexible and fast channel for packaging all of the materials related to one project into a single convenient source familiar to programmers. As I have structured my

workflow around these tools, it has become easier to develop completely portable instruction modules that can be taught or referenced at a moment's notice at any location around the world, with or without internet access (see [PirateBox for Data Literacy](#)).^{Tab 14}

In all of my teaching and workshops about data, I believe it is important to provide the learner with hands-on exposure to real-world examples of data and code. Providing sample scripts to work through realistic and potentially messy data sets is more valuable than a show-and-tell approach with toy examples. Teaching the learner to deal with ambiguity, to expect change, even in the data sources themselves, and giving them tools to resolve the problems that they will face on their own is better preparation in the long run than rote mastery in a controlled environment. After all, in some sense all of statistics is an effort to resolve uncertainty.

Since the use and analysis of data has become a critical skill in today's society across disciplines and levels of study, I have responded by broadening my outreach. In Summer 2016, as part of the *Rutgers Future Scholars* program (futurescholars.rutgers.edu), I began teaching a week-long introduction to data analytics to five selected interns, who are rising juniors in local high schools from low-income families who would be the first in their family to attend college. My philosophy is to expose them to the kinds of questions and software tools that are used by practicing researchers to analyze data. This gives them familiarity and confidence when going forward into college, and demystifies data as a topic for experts only. I taught the Rutgers Futures Scholars again in Summers 2017 and 2018.^{Tab 3, Tab 15} In 2018, I have also taught in Rutgers Executive Education program's Mini-MBA in Data-Driven Management, reaching adult learners.^{Tab 16}

I developed, along with Prof. Javier Cabrera of the Dept. of Statistics and Biostatistics, the Byrne Seminar, "Big Data: Revolution and Reality", which ran for the first time in the Spring of 2016. This course has been fully enrolled at the cap of 20 students every time it has run (Spring 2016, Fall 2016, Fall 2017, and will run again in Fall 2018). It introduces first-year Rutgers students to the impact of big data in various fields through small group discussion and site visits to places where big data is handled in academia (RDI2) and the corporate world (Pfizer).^{Tab 17}

My teaching has also had international impact. In 2017, I delivered a two-day workshop at the *National Statistics Office of Mongolia* on topics in data science with associated R training.^{Tab 18, P3-P6, W3-W6} See these links for an [English](#) and [Mongolian](#) description. Then in 2018, I delivered a one-week "Short Course on Multivariate Statistical Methods with R" at the *Mongolian University of Life Sciences*.^{Tab 18, C3} See these links for an [English](#) and [Mongolian](#) description. I have also been named to the *Fulbright Specialist Roster* through 2020 and am eligible to participate in Fulbright-sponsored host institution projects relating to data and librarianship. I am currently planning these projects with Mongolian colleagues.

Scholarship

My research emerges from my domain expertise, and research opportunities enable me to stretch the boundaries of my knowledge. Refining and sharing expertise via scholarship improves my own practice of librarianship and the discipline as a whole. I have sought to apply advanced methodology and theory from the disciplines of economics and statistics to problems in librarianship, creating scholarship uncommon in the field. I am also committed to open access. Almost all of my publications have been deposited in [SOAR](#), Rutgers' open access repository.

As a business librarian, I wrote and presented about the ways that I sought to advance the field. My *Information Technology and Libraries* article, "Bel Jour: A discipline-specific portal to period-

icals”, described the creation of a subject-specific interactive electronic journal finder at a time when this was a substantial improvement over the discovery tools for e-journals at the time.^{A8} My *Journal of Business and Finance Librarianship* article, “Basic Business Dictionaries Compared”, went beyond a simple review by sampling specific definitions to provide an empirically grounded analysis of the relative merits of the works.^{A6} My second JBFL article, “The Orientation and Training of New Librarians for Business Information” describes the issues involved in introducing core knowledge of business resources to those new to the profession.^{A5}

In addition, my work with BRASS, the Business Reference and Services Section of the Reference and User Services Association (RUSA), a division of the American Library Association (ALA), led to the publication of a chapter on career exploration resources in the RUSA Occasional Paper *Help Wanted: Job and Career Information Resources*^{O3}, and later to my own editing of *Success by the Numbers: Statistics for Business Development*, also published as a RUSA Occasional Paper.^{E1} These comprehensive resource bibliographies exemplify my approach to providing information that is complete and useful.

During this period, my scholarship reflected my disciplinary expertise. I edited the World Wide Web Reviews section in the *Journal of Business and Finance Librarianship* (2006-2011)^{J1-J12} and contributed regular reviews to JBFL and *American Reference Books Annual* (ARBA).^{R1-R14} I also served as subject editor for *Economics of Resources for College Libraries* (2007-2009).

My *Library and Information Science Research* article, “Information Intermediaries and Optimal Information Distribution”^{A7}, was the most thorough application of advanced methodology among my earlier work. By applying economic theory to questions of information production and delivery, I provided a framework through which the appropriate organizational forms for the delivery of different kinds of information could be determined. The insights developed through this approach could be applied to libraries and other information intermediaries to determine their most socially advantageous areas of operation, as well as justify support for specific forms of information distribution. This article has continued to garner growing interest over time, being cited in many disciplines beyond librarianship (Google Scholar 50 citations, Scopus 16 citations), and has been anthologized as an influential paper in *Information Science: Critical Concepts in Media and Cultural Studies, Volume IV, The Economics of Information Science* (Routledge, 2014).

The world of data librarianship is different than the traditional collections-driven world of many other subject specialties in librarianship. Data librarianship is about finding, using, and interpreting data sources, many of which may be freely accessible, but difficult to understand and make use of. In response, rather than working on reviews and editorial responsibilities, I turned my energy to workshops and presentations that would address emerging needs of users for using statistical software, understanding statistical methodologies, and analyzing data. My workshops, as presented at the International Association for Social Science Information Services and Technology (IASSIST) and other venues, emphasize emerging, cutting edge topics. This work then feeds back into my daily practice of librarianship as these leading-edge workshops become adapted to regular presentations to Rutgers students. As mentioned above, the development of these workshops is “demand-driven”, since researchers ask for more coverage of different and advanced topics after initial exposure in workshops.

Beginning with “Introduction to R” workshops at IASSIST in 2010-2011^{W10,W11}, at a time when interest in R software was just beginning to penetrate the library and data landscape, I have developed a series of increasingly advanced and sophisticated workshops for IASSIST, such as “Data Visualization and R” (2013-2014)^{W8,W9}, “Hands-on Big Data” (2015)^{W7}, and “Data Literacy for

All, with R” (2017-2018).^{W1,W2} Selection of IASSIST workshops and presentations is competitive and done by peers in the organization. Presenting workshops for the IASSIST community, itself an organization of sophisticated data professionals, requires me to break new ground and operate at a high level. The value of the content in these workshops becomes clear as they are taken up by other organizations. Subsequently, I have been invited to talk about data visualization at ICPSR’s national conference (2015)^{P11}, for a national webinar series sponsored by RUSA/ALA (2017)^{P2}, as well as at organizations more close to home such as the Institute for Advanced Study in Princeton^{P8} and the Princeton/Trenton chapter of the Special Libraries Association.^{P10}

My national and international presentations also reflect my expertise. Since my first IASSIST presentation (“Doing Data on YouTube”, 2009)^{P20}, I have continued to present on Rutgers’ research data experiences, in outreach in building services (“Designing Data Services for the Institutional Repository”, 2010)^{P19} and training (“Data Management Training to Support Faculty Research Needs: Lessons Learned”, 2012)^{P17}, and with respect to Secure Data (“The Rutgers Secure Data Facility: A Small Scale Solution”, 2013).^{P14} I have supplemented these with regular local presentations, such as on Reproducible Research at the Rutgers Workshop on Reproducibility in Experimental and Computational Science in 2016.^{P9} I was also profiled in the book *Numeric Data Services and Sources for the General Reference Librarian*.¹¹ I presented in 2017 on several topics in Data Science at three major universities in Mongolia, the National University of Mongolia, the Mongolian University of Life Sciences, and the Mongolian University of Science and Technology.^{P3,P4,P5}

Finally, I have written my most recent peer-reviewed articles as a scholar with the perspective of both a librarian and a statistician. I am concerned with how data impacts libraries and how library values such as information literacy and open access are applicable to the world of data. *The Journal of eScience Librarianship* article, “Preparing to Accept Research Data” was a co-authored work that emerged out of the research data team and its planning efforts.^{A3} The *IASSIST Quarterly* article “Data Visualization and Information Literacy”^{A4} explores how data visualization concepts map to information literacy and discusses what data visualization competencies deserve consideration as core competencies for evaluating information in today’s world.

My 2015 article in *PLOS ONE*, “Research Data in Core Journals in Biology, Chemistry, Mathematics, and Physics”^{A2} applied statistical sampling methods to analyze the presence of data sharing in major scientific publications, providing careful and considered evidence of actual disciplinary practice in fields where this had not been closely examined before. By publishing with PLOS as an interdisciplinary open access platform, I also gained visibility outside of library science via tweets and [media coverage](#) in general science outlets.

My *Journal of Academic Librarianship* article, “ARL Libraries and Research: Correlates of Grant Funding” (2016)^{A1} applied a sophisticated regression modeling apparatus to study the question of whether libraries can be demonstrated to have an impact or linkage with the research output of the university. Too often the library literature fails to consider alternative explanations outside of the library field, but this research explicitly includes other research inputs, which in this case are more closely related to research outputs. This article was also distributed in arXiv in support of interdisciplinary open access.^{O1} I have exemplified the practice of open data reproducible research in these most recent articles by placing my data and software code in openICPSR so that they can be freely accessed and examined.

Service

My primary goals in service have been to advance the organizations that support my work and the development of my specialties in business and data. I have sought and attained leadership roles in those organizations. In addition, I have sought to make positive contributions to the University.

While I served as a business librarian, my primary professional commitment was to BRASS, the Business Reference and Services Section of the Reference and User Services Association (RUSA), a division of the American Library Association (ALA). During more than a decade of activity in BRASS, I rose to the position of *Chair of BRASS* in 2006-2007. I served as chair of several BRASS committees and BRASS Representative for others. One highlight was chairing the Program Planning Committee for the annual BRASS program in 2004. The Chair of BRASS role entailed several years of participation in leadership duties in addition to the term as Chair, including strategic planning for the future of the organization.

When my role changed to data librarian, I joined IASSIST, the International Association for Social Science Information Services and Technology. IASSIST (iassistdata.org) is an international organization that brings together data professionals from around the world who work in libraries, statistical agencies, and data archives. In this community, I found support and learning opportunities for my data interests, and I again sought to serve and advance the organization, serving on the Program Planning Committee in 2010, a survey team supporting the IASSIST Strategic Plan in 2011, and several interest groups. In 2015, I was elected to a four-year term as *IASSIST Secretary* (2015-2019). The Secretary is one of four members of the Executive Committee of IASSIST, and as such, participates in all leadership decisions and the larger Administrative Committee's monthly meetings. In addition to taking minutes, running elections, and other record-keeping, the Secretary performs a variety of tasks supporting the effective operation of the organization, including strategic planning, and backing up the Treasurer and President when necessary. As Secretary, I drafted IASSIST's [statement in response to the President's executive order on visas and immigration](#) in February 2017. It has been a privilege to be trusted as one of the key people sustaining the functioning of IASSIST and its now over 500 members during a time of growth for the organization.^{Tab19}

At the national level, I served on the *ICPSR-IMLS Project Advisory Committee* (2010-2012), which explored social science datasets with the intention of developing best practices and services for institutional repositories. I also served on the *TeachingWithData.org Advisory Group* (2009-2011) to advise this NSF-National Science Digital Library project supporting quantitative literacy in education. These projects deepened and extended my knowledge of data librarianship. At the regional level, I have served as a member of the New Jersey State Data Center Advisory Board since 2005.

In support of Rutgers University, I have served on search committees and task forces such as the Committee on Work and Family Issues (2008-2009). Most recently, I was elected to the *New Brunswick Faculty Council* for a 2017-2020 term, and was appointed as *co-chair of the Council's Committee on Budget, Planning, and Infrastructure*. I am committed to understanding and improving Rutgers University as a whole.

While some of the service roles directly related to my library duties have been discussed above, I have also served the Rutgers University Libraries system and the New Brunswick Libraries through more general roles, including on search committees and as a mentor. I will

highlight a few that represent particular leadership. As *Faculty Coordinator* for the Libraries in 2012-2013, I led investigation into the future of research services and future librarian skills. Based on surveys and discussion sessions involving all library faculty, these initiatives culminated in two detailed reports summarizing the consensus of the Libraries at the time.^{Tab 20} From 2007 to 2008, I *chaired the New Brunswick Libraries Faculty* and represented NBL in systemwide bodies. Another initiative reflecting my interest in international outreach was my time as *Chair of the World Languages Task Force* (2006-2011), during which I led a team that developed translations of a welcoming message from the Libraries to speakers of ten languages, including Arabic, Chinese, Hindi, and Russian.^{Tab 21}

Conclusion

My librarianship, scholarship, and service are closely linked to my ideal of a data librarian as an expert in methods and resources who is also focused on education in data and statistical literacy, functioning on a global stage. When my primary duties were in business librarianship, I published and served as appropriate for my specialty (JBFL, BRASS). During my last decade as data librarian, I have become a leader in serving the field (IASSIST Secretary), in publishing about data using advanced statistical methods, and in leading workshops to share data and statistical knowledge at all levels from local New Jersey high school students, through Rutgers students and faculty, and beyond to the world, via open access web materials and outreach to Mongolia.

My workshops and teaching exemplify my approach to empowering users with a hands-on approach to real data and sophisticated tools. My recent scholarship represents what I stand for, as a data specialist informed by library values and statistical expertise, exploring the changing scholarly landscape while being committed to open data and open access. In addition, I exercise leadership in advancing Rutgers University, the profession, and the scholarly community as a whole.

My professional goals are to continue along the path I have laid down, developing further mastery of advanced statistical methods and specialized data resources, then bringing that knowledge to the researcher through multiple modes of publication and presentation. This is an adaptation of the librarian's traditional mission to help users find, access, and interpret the information they need to answer their research questions, only now the tools are cloud-based R software, Github, online open data repositories, and YouTube. One may say, "plus ça change..." but constant adaptation to the most recent developments, scholarly, technological and otherwise, is what is necessary to make new discoveries about the forest while others are examining the same old trees.