

What Do Data Librarians Think of the MLIS? Professionals' Perceptions of Knowledge Transfer, Trends, and Challenges

Camille V.L. Thomas and Richard J. Urban

There are existing studies on data curation programs in library science education and studies on data services in libraries. However, there is not much insight into how educational programs have prepared data professionals for practice. This study asked 105 practicing professionals how well they thought their education prepared them for professional experience. It also asked supervisors about their perceptions of how well employees performed. After analyzing the results, the investigators of this study found that changing the educational model may lead to improvements in future library data services.

Introduction

Data services are growing in academic libraries. Services such as data management, data curation, and data visualization are parts of the larger research data lifecycle. With the advancement of technology, scholars can make use of data to produce new kinds of scholarship. The need for research data services is due to the growing amount of data available, federal funder data management mandates, data-sharing networks, and assessment based on data analysis. Change to the data lifecycle reveals an opportunity for libraries to support researchers. Since 2011, several library science master's programs began offering courses and certificates concerning the research data lifecycle. Data Curation specializations within MLIS programs are designed to provide employable skills for the crucial, growing need for data support services. However, the skills and topics covered in Master of Library and Information Science (MLIS) programs are not adequate for practice in the field. Previous research only examined the adequacy of education programs according to program website content, course descriptions, and job descriptions. This study will draw from the opinions of practicing professionals to fill gaps in the literature and offer effective solutions to prepare students for data services in the future.

Camille V.L. Thomas is Scholarly Publishing Librarian at Texas Tech University Libraries; e-mail: camille.thomas@ttu.edu. Richard J. Urban is Digital Asset Manager and Strategist at Corning Museum of Glass; e-mail: richardjurban@gmail.com. Preliminary findings were presented at a poster session at the 2015 National Conference of African American Librarians (August 6, 2015, in St. Louis, Missouri), which was included on page 239 in the conference proceedings. ©2018 Camille V.L. Thomas and Richard J. Urban, Attribution-NonCommercial (<http://creativecommons.org/licenses/by-nc/4.0/>) CC BY-NC.

To give a clear background for this emerging skill set, the investigators of this study will discuss: definitions, the development of data curation, data science in other fields, data services, and data curation education. It is important to note that definitions of data curation still vary. The definition of data is becoming broader, as digital humanities and interdisciplinary digital scholarship develop. When defining “digital curation” and “data curation,” the library and information education curricula make only minor distinctions.¹ Many definitions limit data curation to research in hard sciences, business, and social science research. Data is associated and valued in the sciences because replication and validation are basic principles of the scientific method.² Making data easier to replicate and contextualize helps researchers confirm or deny scientific claims. However, “digital curation” was associated with text or images, primarily used in the humanities. “Digital curation” is broader, defined as “the active management and enhancement of digital information assets for current and future use.”³ Data can take the form of text, images, video, spatial data, computational code, and more. Data curation refers to organizing, maintaining, integrating, and extracting knowledge from data to apply and reuse it.⁴ In this study, we will include several different data types when we use the term “data curation.”

Data management and data literacy are common services in research libraries, but they are only two aspects of research data lifecycle. Data literacy is the ability to “read, create and communicate data as information.”⁵ Data literacy instruction is often the first step in supporting researchers, encouraging updated data practices, and growing data services.⁶ The difference between management and curation is that “data management” emphasizes organizing, collecting, describing, and storing data for communication among scholars.⁷ “Data curation” includes “systematic and purposeful” extracting for general purposes such as public use.⁸ In other words, data curation is like curating a museum collection for exhibit rather than for internal storage. Curation is an important aspect of the lifecycle that can only occur when data management has been properly planned.⁹ Data science refers to the analysis, visualization, or interpretation of data.¹⁰ There should a distinction made from “digital preservation,” which is defined as “the series of managed activities necessary to ensure continued access to digital information for as long as necessary,” focused on archiving.¹¹ Digital preservation emerged during the late 1990s to address new media’s vulnerability to deterioration and catastrophic loss after less time than traditional formats such as paper, microform, and photos.¹² Data preservation is an extension of digital preservation, as digital data degrade quickly and require frequent validation checks, remote backups, and metadata prior to storage.¹³ Again, distinctions vary among different organizations. In this study, management and preservation are considered processes included in curation. In the context of this study, “data curation” refers to “management activities required to maintain research data long-term such that it is available for reuse and preservation.”¹⁴

It is important to know how effectively data curation knowledge is transferred from education to professional practice as data curation programs develop. The investigators used the following questions to guide the study:

- RQ1: Do librarians providing data services believe their education prepared them for the job?
 - RQ1A: Did they attend data curation programs or MLIS programs without data courses?
 - RQ1B: Do they have any other education or experience that prepared them for the job?
- RQ2: What knowledge areas or skills do they use on the job?
 - RQ2A: What knowledge areas do they lack that might prepare them for data services in the future?

- RQ2B: What advantages, challenges, and trends do they see as the job develops?
- RQ2C: How well do supervisors believe their employees' skills have prepared them for the job?

The purpose of this study is to understand professional perceptions of effectiveness. The intended outcome is to use the data to identify areas for improvement in the future development of the workforce and services in information organizations.

Literature Review

Data curation is a growing area of employment need in both private and public sectors. Federal and private funding agency mandates for data management, open publication, and data sharing have compelled scholars to reexamine their practices. Agency policies grew from a few in 2011 to 15 as of 2015.¹⁵

In order for data to be fully and publicly accessible to search, retrieve, and analyze, specialized curatorial actions should be taken to prepare the data for reuse, including quality assurance, file integrity checks, documentation review, metadata creation for discoverability, file transformations into archival formats, and selection of a suitable license/ copyright.¹⁶

Since universities produce the majority of research materials, it is imperative for academic libraries to offer support.

Libraries are evolving to address new data policies and practices. Training, professional task forces, and federal policy began to develop between 2006 and 2011.¹⁷ A study in 2010 found common job titles and descriptions required an LIS education with an emphasis on metadata creation and data management in approximately two-thirds of job advertisements.¹⁸ The Association of Research Libraries' (ARL) Data Curation SPEC Kit found that data curation services appear to be a relatively recent initiative; more than half of the libraries that currently provide services (35 of 51) started doing so in 2010 or later.¹⁹ In 2013, *Library Journal* began to recognize a "small subset" of job placements in data curation as well as a decrease in reference jobs since 2011.²⁰ Recent placement reports identify seven librarians placed in data analyst jobs in 2015 and five librarians placed in data curation jobs in 2016.²¹ Emphasis on skills for handling data is evidence of the growing need for data professionals.²² Outside the field, librarianship is still considered traditional technical labor (that is to say, cataloging), excluding emerging specializations such as "data librarianship." *Preparing the Workforce for Digital Curation* reported no growth in "digital curator" job openings in the library field in the past seven years.²³ In contrast, Si et al. found there had been growth in education and hiring for data services.²⁴ It may seem like a challenge for library science to keep up with developments in other industries such as business. However, the vast scope of the work could create an opportunity for collaboration and will not be "isolated in libraries or in any one type of institution or organization."²⁵ We are still defining, developing, and promoting our roles as stakeholders in the research data lifecycle.

Since 2010, those who were hired or appointed to data librarian roles provided fundamental services such as data management and data literacy. A study by Tenopir et al. in 2015 gathered 302 responses about the perception of data services from a stratified sampling of librarians and library directors.²⁶ Of those surveyed, 83 percent of libraries offered reference services for finding and citing data sets or planned to in the next two years, suggesting libraries also attached great importance to the application of skill sets related to data. As of 2017, 45 libraries (92%) provide one or more curation services, and all but chain of custody are offered by more than two-thirds of

the libraries.²⁷ Hands-on experience with data is needed so professionals are able to work with researchers throughout the research process.

Communication with researchers is important because, “although processes, institutions, and skill sets for preserving and disseminating digital information have been known and in place in some disciplines for several decades, identifying that assemblage of practices is essential to fully establish the field of digital curation.”²⁸ Results show that libraries commonly offer reference support for finding and citing data. However, outreach and collaboration with partners on data curation were the least commonly offered services.²⁹ There is an emphasis on the sciences, which has the most growth and demand for curation services, in everything from academic STEM research to K–12 education.³⁰ Therefore, services are still limited to data-driven subjects and are not offered as a suite of comprehensive services available to all disciplines.

As we seek to develop our roles, services can mature from fundamental data management services to services like data curation. Librarians have the collecting, selecting, and organizing skills to carve a role for themselves as data curators. Ogburn notes that libraries have collected social science, census data, faculty papers, and family records in physical and electronic formats for a long time.³¹ Traditionally, the acquisition, stewardship, and preservation processes for these materials was straightforward. Librarians must now adjust information curation skills to reflect the digital formats that are often not well organized by scholars.³² These skills, which are long-standing in the profession, just need to apply in a new context. Changing data practices means a fundamental change to the research process. Therefore, the structure of library support and librarians’ skill sets must change because data curation requires cooperation to sustain collaborations that “transcend disciplines, departments, service units, and institutions.”³³ Librarians are skilled in working across many different disciplines, allowing them to mediate practices that are often siloed and software-intensive. Weber emphasizes human interaction with knowledge infrastructure as the “key to developing and sustaining a global system of interoperable digital data and tools” across disciplines and organizations.³⁴ We can take on the role of adding context and provenance to data sets by ensuring proper compliance, quality, appraisal, and design. If librarians take on the role of data curator, time-consuming tasks become more manageable for users. In this emerging role, libraries can act as curators, facilitators of curation, or both. To meet researcher needs and fulfill a new role for libraries, we need librarians with the appropriate skills.

There is often miscommunication between libraries and data librarian candidates about which skills are needed for open positions. Miscommunication occurs during the hiring process. Job titles in the field lack consistency (such as Data Management Librarian, Data Curator, Data Librarian, Data Services Specialist, and E-Science Librarian) because previous knowledge was based on theory in education programs rather than the practice of library services.³⁵ Job descriptions often include all possible competencies in a “wish list,” rather than an accurate account of the skills needed for the particular position and organization.³⁶ Outside the library science field, Aaron Kimball, Chief Technology Officer of Wibidata, an analytics software company, expressed concerns about this problem as well. He feared that the data scientist role would become like the “mythical” webmaster of the 1990s.³⁷ Koltay observed that the library science field did not agree on a single definition or vocabulary as of 2015.³⁸ Kitchen-sink job descriptions reinforce the lack of consensus on the role of libraries in data practices discussed in the study by Si et al. Now that libraries are expanding their roles in the data research lifecycle, it is important to make distinctions between responsibilities going forward.

The lag in libraries about definitions, skills, and responsibilities may have already affected the workforce. Interest and growth between 2010 and 2017 are evident, but true

investment in these services is lacking. For example, 35 percent of 2,162 alumni from 41 LIS programs reported placement in a data curation job or responsibilities in 2011, but placement reports from 2012 through 2016 do not match those findings.³⁹ *Library Journal* placement statistics report a drop in data curation placements between 2013 and 2014 (10 down to 4), then again between 2015 and 2016 (7 down to 5).⁴⁰ However, in 2012, *Library Journal* did not recognize the “data curation” title yet. Jobs categorized under “metadata,” “digital library,” or “records management” might have started to include data service responsibilities. The fact that “digital curator” labor is not represented in the Bureau of Labor Statistics Standard Occupation Classification suggests many of these duties are still referred to as “records management.”⁴¹ Additionally, the development of specialized research data services is often constrained by knowledge and skill gaps among library staff, unclear expectations, and lack of confidence in their expected roles in research data services.⁴² Libraries should demonstrate reasonable expectations for candidates to perform fundamental services, along with plans for long-term investment in the maturation of data services. Distinguishing between goals and immediate expectations could solve the problem of attracting candidates with core competencies and communicating the belief that they will build their confidence, skills, and services on the job.

To get to the root of the issue, we examined competencies for data services and how librarians are trained. Organizations including the International Organization for Standardization, National Institute of Standards and Technology, and the Digital Curation Centre in the United Kingdom worked to determine competencies. The work of those organizations makes the need for and development of a standardization framework, as well as core competencies in data curation education, evident.⁴³ Goodsett and Koziura reviewed shifts from a traditional library curriculum to an information science curriculum circa 2005, which emphasized data and technology.⁴⁴ Palmer noted increased discussions of best practices for data curation in Library and Information Science (LIS) programs and continuing education since 2010.⁴⁵ Varvel, Elin, and Palmer conducted a study of various programs in 2012, revealing the curriculum in LIS education for data professionals had “been attached great importance.”⁴⁶ MLIS programs then sought to instill competency and confidence in the next generation of library and information professionals through the development of data curation curricula. Varvel, Elin, and Palmer believe librarians often do not feel confident in their abilities because theory is taught in MLIS programs instead of training for practice.⁴⁷ Their findings match Goodsett and Koziura’s 2016 study, which gathered perceptions of recent graduates in many different jobs to identify the gaps in general LIS education.⁴⁸ Just as library services evolved to address data curation needs, Library and Information Science education programs also sought to address the growing need for librarians to acquire appropriate skills.

Many studies within the literature focus on how data curation preparedness looks on paper (in other words, job descriptions, course offerings, LIS program websites). One study by Si et al. and another by Harris-Pierce and Liu highlight educational institutions that launched pilot programs.⁴⁹ These programs include the University of Illinois, the University of North Carolina, and the University of Arizona with grants from the Institute of Museum and Library Services (IMLS), as well as University of Michigan’s “Preservation of Information,” a specialization with a digital curation focus. Online course catalogs and program websites of 63 iSchools and other LIS schools divided these courses into four categories: data-centric, data-inclusive, digital, and traditional LIS.⁵⁰ A few programs such as the University of Urbana-Champaign, University of Pittsburgh, and Carnegie Mellon University Libraries used partnerships with data scientists or data centers to offer practical application within their education programs,

subsequently showcasing librarians as stakeholders in the management of data.⁵¹ These studies tell us about existing programs, competencies identified, and models used to teach them. However, there is very little scholarship on perceptions of data education adequacy when put into practice.

A report from the University of Illinois gave insight beyond MLIS program websites or course catalogs. This report shows results from an informal study to gather the perception of job preparedness from alumni who completed the data curation specialization.⁵² It is limited in that it only collects information from University of Illinois alumni. The alumni are currently working in data curation positions or have other jobs that require them to perform related duties. Of those working in a data curation position ($n = 10$), six graduates were working in academic settings and one each were in government, research center, and data center. The survey includes response results about courses, job placement, institution type, job responsibilities, and desired continuing education. It creates more defined duties by collecting information from a variety of employers based on alumni responses. The most frequently reported were positions in an academic setting (49%), corporate (17%), and nonprofit (15%) settings. Government agencies hired five graduates, research centers hired three graduates, and data centers hired two graduates.⁵³

University of Illinois alumni desired further education in data interpretation, research administration, and proposal writing. Goodsett and Koziura's general study identified research methods as the most valued experience of new librarians' education and the largest gap in their knowledge.⁵⁴ Luo conducted a study in which practitioners in many areas of librarianship wished they had learned more about research methods over a longer time period, including statistics.⁵⁵ In that 2011 study, the majority (55%) received their MLIS more than seven years ago and 31 percent did not take a research methods course. The types of continuing education desired by University of Illinois alumni reflects suggestions from Webber as well as Luo, that MLIS programs need state-of-the-art computational skills such as statistical analysis. All the studies so far either have a narrow focus on specific programs or a wide focus on a range of librarians with very different skill sets. It is important to consider data education from other fields to pinpoint areas of competency because librarians have such varied educational backgrounds that can be used.

Programs in fields other than Library and Information Science emphasize computational skills more frequently. Many programs in other fields offer matriculated graduate degrees and certificates in "data science," "machine learning," "big data," and "analytics." Coined by D.J. Patil and Jeff Hammerbacher in 2008, "data science" is the most commonly used term in other fields.⁵⁶ Again, data science refers to the engineering, analysis, and interpretation of data. Electrical engineering, business management, statistics, and computational science programs usually contain data science programs, but interdisciplinary programs also contained data science programs as of 2013.⁵⁷ Song and Zhu assert that people who have training in computer science, statistics, and mathematics are considered data scientists. These data scientists are often salaried at twice the rate of data librarians.⁵⁸ Typical courses required by data science bachelor's programs include Probability and Statistics, Data Mining, Discrete Mathematics, Data Structures and Algorithms, Database and Machine Learning. However, information science programs also offer data programs because of their emphasis on human-computer interaction. Popular courses in data science graduate programs deviate from a quantitative curriculum and include courses one might find in an iSchool such as Text Mining, Information and Social Network Analysis, and Data Visualization.

Competencies for data scientists and data librarians are similar: the shape of data infrastructure and its semantics, data discovery, visualization, processing, infrastruc-

ture. Demchenko and Los include capabilities for data analysis, uses of data and its implications, curation, as well as archiving, preservation, storage, permanency, data (de)selection, and annotation. They also list project management, leadership, safety, financial, and legal considerations as competencies.⁵⁹ Song and Zhu identify the 6-step Cross Industry Standard Process for Data Mining (CRISP-DM) model as the most prevalent in data science education. However, Howe and Palmer assert that there is a spectrum of useful competencies, ranging from programming, which centers on infrastructure, to analysis, which focuses on abstract concepts.⁶⁰ Even Song and Zhu, who view only people with computational skills as data scientists, recognize that cultures and strengths differ among institutions and conclude that the biggest hurdle is allowing time for the development of capable data scientists.⁶¹ Los and Demchenko believe data science professionals will work in research, specific scientific data domains, and the private sector. Scholars, especially those in the sciences, are concerned with competencies in varying data types, data storage, and management of data. These are the same types of institutions that hired University of Illinois alumni. These are the same competencies with which data librarians are concerned. Comparatively, libraries are not too far behind data science regarding exploring the new frontier.

Various perceptions from information professionals in other fields can be insightful to librarians who may not gain confidence from the MLIS. Charles Pensig, director of data science at Oak Labs, Inc., wrote a post on why he dropped out of a data science graduate program.⁶² He explains that data science programs are still exploring an emerging area compiled from long-standing practices in existing fields but in a new context. Institutional education means more accountability, networking, and signaling to potential investors or employers. Pensig counters that high-functioning individuals are likely to learn those competencies on the job or be self-taught through online resources or books. While Pensig's perspective is only one example, it shows that education studies in other fields also give incomplete, inapplicable information and may not take into consideration the perceptions of those doing everyday work.

A common thread in library data services is the need to inform and engage scientists and scholars who often perceive they do not need to engage with libraries.⁶³ However, *Preparing the Workforce for Digital Curation* cited automation as the major topic for future development among data scientists. Slow development in education and the desire for automated processes show that researchers working with data do need help from knowledgeable counterparts. Los and Demchenko's curriculum strategy includes collaboration between data scientists and academic programs much like the model at the University of Illinois Urbana-Champaign.⁶⁴ There are already at least 50 programs in "data science." There are 11 data programs out of 63 library science or information schools with accreditation from the American Library Association.⁶⁵ If curriculum in LIS develops with competencies equal to those of programs in other fields, the profession may be able to fill existing gaps better.

To summarize, the existing library literature primarily focuses on objectives in job or course descriptions. This is mostly speculation based on what library services and education hope to achieve. The library profession, along with many other professions, recognizes the growth and importance of curating big data as the volume of data increases. However, discussions are usually limited to the echo chamber of each profession. The problem is this means library and information education programs have not revised their curriculum to keep up with competencies needed in practice or to keep up with data curation in other professions. If library professionals do not agree on one definition of what data curation means and do not address the gap between theory and practice, there will continue to be a miscommunication with researchers. Yes, data curation may seem like a new and complex specialization for libraries to undertake,

but many librarians are already qualified and practicing. This study sought to address the gaps by asking data curation professionals specifically what connections they draw between their educations and their experience on the job.

Methodology

We created the survey using Qualtrics as the platform. Investigators anticipated inconsistencies in titles, self-identification as a “data librarian,” and ad hoc data responsibilities. There was no confirmed, finite number of professionals performing data curation services. However, in 2014, Tenopir et al. collected a total of 302 responses from librarians and directors at 223 research-intensive institutions. We did not collect data on ethnicity, gender, or age, as it was not relevant for the scope of this particular study. We did not collect data about when librarians attended MLIS programs or length of work experience. We anticipated responses would come primarily from research universities, particularly in the United States, but results categorized by institution were also unmeasured at the time of the study.

The investigators sought a broad sample of institutions. We did not want to limit the sample size to universities with iSchools featuring data curation programs to avoid confirmation bias. The investigators of this study disseminated the survey via industry e-mail listservs including those focused on assessment, diversity, geoscience information systems, political and social sciences, scholarly communication, and government documents. Informal professional networks were also used, specifically through word of mouth, blogging communities, and social media (Databrarians, #databrarians, #data, and the like). We chose these channels because they were recommended by professionals in the pilot group to be the most effective ways to reach as many librarians who worked with data or data-driven disciplines as possible.

The survey gathered anonymous information about workplace types, educational background, the perception of courses taken, perceptions of preparedness through education or practice and desired continuing education. The survey design included multiple-choice, Likert scale, and open-ended questions. Online surveys are able to reach more people at a faster rate, so open-ended questions were included in place of interviewing. Investigators coded and analyzed qualitative data using grounded theory, co-occurrence and sentiment analysis. Some qualitative data was summarized as context for quantitative results. Variables of interest include the type of education, whether education included data curation-specific coursework, satisfaction with education, the perception of the primary source of skills learned, the perception of education still needed to perform competently, and perception of future developments within the specialization.

Many of the questions regarding knowledge or skill area were based on the survey the University of Illinois' School of Information distributed to its alumni. We also used course catalogs from data curation or digital curation programs to develop questions related to courses. We did this to gather information about programs that do not offer specializations in data curation but may provide relevant skill sets. We also wanted to add more depth to the existing information about preparedness, which centers on course offerings.

Although the primary investigators were limited to indirect practice, they were knowledgeable about the nature of the work. Thus, the investigators clarified questions by piloting them to avoid bias. The goal was to design the survey in as much context as possible. Pilot responses were taken into account but not analyzed with the rest of the responses since the data was not anonymous. Spambots accounted for approximately 115 of the 220 total original responses, with some margin for those abandoned by humans. Surveys less than 20 percent complete were discarded, leaving

105 responses used in the analysis. This is still a representative group when compared to Tenopir’s sample of 302 librarians and directors combined, but again there was no record of the total number of data librarians at the time of the survey.

The survey gathered nonprobability sampling. Ethical considerations for this proposal included voluntary participation, documented by a consent agreement to avoid harm to participants. Likewise, we informed participants of the anonymity and confidentiality of the survey. The investigators hold ethical obligations to the International Review Board, as well as professional standards for analysis and reporting. We did not alter results in any way. To increase the impact of the survey and encourage further research on the topic, results were intended to be disseminated widely and publicly.

Results and Analysis

We focused on attitudes about the realities of daily work and education rather speculation based on professional or learning objectives. The investigators desired responses from any professionals working in a capacity in which they provided a service that supported data created by researchers. Any and all experiences were necessary to address gaps in the existing literature, educational models, and hiring practices.

The majority of participants came from North America, but there were some results from European and Australian participants. There were 105 total respondents, 70.48 percent of whom worked in an academic library, as hypothesized. Answers such as “Research University, nonlibrary” and data centers (data archives or data repositories) made up 9.52 percent “other” responses. Professionals then reported their educational backgrounds.

A few respondents reported earning another master’s degree or PhD in addition to the MLIS. However, in another portion of the survey, other respondents reported they neither had library education nor worked in a library. Of the responses (n = 102), 70.41 percent earned their MLIS degree, 5.10 percent earned some MLIS credits, 6.12 percent earned a specialist degree, 1.02 percent earned a doctorate degree, and 21.43 percent had no formal education from a library program.

We hypothesized that most respondents would attend universities that offered data curation courses even if they did not take those courses. We thought those universities were on the cutting edge, perhaps teaching relevant knowledge through other outlets. The investigators wanted to build on previous studies that focused only on

**TABLE 1
Professionals’ Institutions of
Employment**

Institution Type	Percent of Professionals (n = 105)
Academic Library	71%
Research Center	6%
Private Sector	2%
Public Library	3%
Community College/ College Library	2%
Government Agency	2%
Nonprofit Organization	4%
Other	10%

**TABLE 2
Highest Degree Earned**

Degree	Percent of Responses (n = 105)
High School Diploma or GED	1%
Associate’s Degrees	0%
Bachelor’s Degree	26%
Master’s Degree	47%
Specialist’s Degree	1%
Doctorate Degree	22%
Other	3%

programs with data curation specializations. Results revealed most professionals (74%, n = 85) attended universities other than the institutions listed, which had MLIS programs with current courses relevant to data curation. One professional noted having more than one degree from multiple institutions we listed. The survey did not ask about MLIS programs or accreditation from the American Library Association in this question.

TABLE 3
University Attended (MLIS Programs with Data Curation Specializations)

University Attended (US)	Number of Responses	Percent (n = 85)
University of Illinois–Champaign-Urbana	8	9%
University of Maryland	0	0%
Rutgers University	2	2%
University of North Carolina–Chapel Hill	2	2%
University of Arizona	1	1%
University of Michigan	4	5%
Simmons College	5	6%
Other	63	74%

As far as specific data curation education, 91.95 percent (n = 89) did not have specialized certificates, concentrations, or formal training in data curation. Only 9.20 percent reported any formal data curation education. We hypothesized that a direct connection would exist between information professionals who had training in MLIS data curation programs and those who were hired in data curation roles within libraries. The investigators did not expect most professionals to have data curation specialization, but we did expect that the number of those who specialized in related knowledge areas to be higher.

The top knowledge areas that practicing professionals took in their coursework included Information Organization, Database Management, Metadata Theory, and Digital Libraries (n = 77). Courses input by respondents included Data Mining, Statistics, Intellectual Property Law, and Web Development. Also, respondents listed degrees in other data-driven disciplines such as social sciences, geoscience, systems engineering, and health informatics.

Additionally, in the qualitative responses, many participants said their MLIS program did not offer several courses specific to data curation at the time of their attendance. Low response rates for knowledge areas such as Information Storage and Retrieval, Systems Analysis, Digital Preservation, and Information Modeling reflected those statements.

Professionals spent a large portion of their work consulting on data management policy, performing data reference and literacy instruction rather than performing curation, as in previous studies. However, they still considered backgrounds in data-driven areas valuable for understanding how to manage data and work with researchers. Participants worked with many different types of data. Further clarification for the other field included: “biomedical,” “time series,” “interactive digital humanities elements,” “instrument,” and “physical specimens” data types.

Perceptions of how well the MLIS prepared professionals for their work was divided. Sentiments were mostly neutral, slightly leaning more toward positive perceptions of

TABLE 4
Data Types Handled by Professionals

Data Type	Affirmative Responses	Percent (n =104)
Text	102	98%
Images	93	89%
Presentations	72	69%
Videos	80	77%
Spreadsheets	91	88%
Databases	83	80%
Spatial Data	66	63%
Computational Models	50	48%
Computational Code	61	59%
Raw Data Sets	78	75%
Other	9	9%

preparedness. On a scale of very well to poor, 33.79 percent (n = 74) responded positively (Well, Very Well), 28.38 percent responded neutrally, 29.73 percent responded negatively (fairly, poorly), and 8.11 percent did not attend an MLIS program.

Open-ended responses gave richer insight into the reasons behind their perceptions. For example, several learned about data curation in other roles. The sentiment was split, reflecting the quantitative data. There were 27 negative qualitative responses and 15 positive responses when asked about how the MLIS prepared them for professional work in data curation, in which 20 respondents had an MLIS and seven did not or

attributed preparedness to education in other disciplines. For example, one professional wrote:

“It has, but not traditionally. I am a research scientist who switched fields and now works as the binding link between researchers and our digital curation office. That way, I know the needs of the research community and can communicate those easier to the library services and vice versa. A lack of communication or even misunderstanding are the main obstacles for an effective digital curation.”

However, another responded: “I actually like the way my school [University of Illinois Urbana- Champaign] had it set up.”

Also, negative sentiments were frequently associated with mentions of education when coding responses. Respondents also seemed to associate education with “outdated” or “underdeveloped.” The following response is an example:

“At the time I was in my LIS program, the courses in data curation were really just beginning. I took one that focused on scientific data curation that was more on the theory/philosophy and less about actual practice (i.e., what it takes to curate a data set).”

Professionals often felt the survey was biased toward newly graduated librarians or information professionals who may have had more updated training or coursework. Knowledge areas listed were based on course offerings in library programs’ data curation concentrations, which are more recent. However, the investigators of this study made no indications of value for which knowledge areas were relevant. Furthermore, those who received LIS data curation coursework in recent years voiced similar concerns. MLIS data curation alumni perceived their education as too theoretical and, subsequently, had to learn a great deal on the job. Although the survey did not collect years of attendance, those with “updated” education and those with “outdated” education had the same perceptions of adequacy.

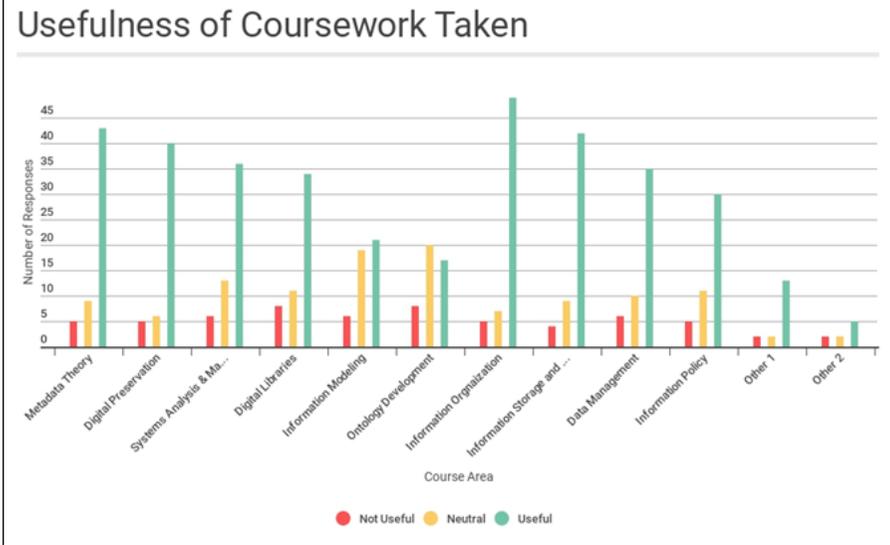
TABLE 5
Learning Channels for Knowledge Area

	Learned via Coursework	Learned Equally via Coursework and Hands-On	Learned via Hands-On Experience	N/A or No Coursework in This Area	Total Responses
Metadata Theory	11	19	19	28	77
Digital Preservation	13	8	18	38	77
Systems Analysis and Management	9	12	17	37	75
Digital Libraries	2	19	20	36	77
Information Modeling	6	8	15	45	74
Ontology Development	10	6	11	45	72
Information Organization	9	34	16	17	76
Information Storage and Retrieval	8	21	19	27	75
Other 1	3	5	4	9	21
Other 2	1	2	2	8	13
Data Management	1	19	29	26	75
Information Policy	9	14	16	34	73

The most useful courses taken included Metadata Theory, Digital Preservation, Information Organization, Information Storage and Retrieval, Information Policy, and Data Management (n = 74). Perceptions of usefulness varied. Alternative responses in the “other” field included Scholarly Communication, Copyright/Intellectual Property Law, Records Management, Information Systems Management, and Information Literacy Instruction. Many of the professionals did not take recent course offerings, so the knowledge areas we expected to have high ratings for usefulness did not have a high rating in the results. However, several knowledge areas (policy, applied research methods, records management and managing relationships with researchers) were considered useful. It was difficult to tell if professionals considered some knowledge areas useful for practical application because of the high rate of neutral responses.

This may be why there is a lack of consensus about a definition of data curation and of job descriptions in the field. Many other “kitchen sink” job descriptions also contain an unmanageable list of skills and/or duties for one professional. Data jobs could be asking for “unicorn” job seekers, inadvertently becoming unattractive to professionals who are qualified. It may improve services if qualifications and responsibilities are prioritized and concise. Professionals could then learn about secondary responsibilities during the hiring process or expand the role after hire. It would also be insightful to emphasize analytic and human-computer interaction skills.

FIGURE 1
Ratings for Coursework/Knowledge Area Taken



Professionals also suggested courses for MLIS programs to adopt. In summary, participants of the survey suggested applied data courses such as Research Methods (often expressing students should learn methodology practices across disciplines), Data Analysis and Computing, and Statistical Assessment. For example:

“A course on how scientists (humanists/social scientists, not librarians) think and work and communicate, workflows and workflow tools, emerging technologies for creating, managing, sharing digital scholarship. A basic course in R or other statistics tools, bibliometrics.”

Professionals also suggested policy and relational courses such as Scholarly Communication, Data Management, and Personnel Management.

Those in supervisory roles made up 30.77 percent supervised staff who worked in data curation roles (n = 92). Respondents reported an average of 7 FTE employees dedicated to data curation at their institution and a median of 1.25 FTE. Several respondents had difficulty with this measure due to varying types of employees and definitions of relevant work at their institutions (n = 26).

The majority of those staff earned a master’s degree as their highest level of education. Ninety percent of supervisors reported that their staff did not have data curation specialization as a part of their education, while 13.64 percent reported that their subordinates did have formal data curation education. Supervisors then reported certification for staff, whether formal or informal. The majority of staff had no certification, but many had professional development or continuing education.

Supervisors rated the overall adequacy of staff performance (n = 25). The majority of responses were “not applicable” or “unsure” (43.48%), which may have been due to reluctance to give negative results or due to a lukewarm perception.

Finally, 21 percent stated that their staff did not attend an LIS program, showing that any library training was on the job. However, when asked if staff education had equipped them for their job in data curation, 52.63 percent said Yes and 47.37 percent

said No (n = 20). They were also asked to give a brief explanation for their answer. Many stated that staff learned hands-on skills under their supervision, while others attributed foundational skills for their job performance to their education, especially those staff with library education.

Professionals reported perceived advantages and challenges at their current institutions to identify any gaps in on-the-job training. The most common responses for both challenges and advantages were related to institutional support. Other responses included relational issues such as silos and politics among institutional library or center departments, discipline-specific practices, and communication. There were also technical and infrastructural concerns about personnel, insufficient technology, lack of robust storage, and the ability to expand their skills. The desire for automation showed up frequently in the qualitative responses but not in the quantitative data. Time and time-consuming processes or workflows were a common challenge. Also, some responses specified that not only was funding needed but separate long-term and short-term funding.

Challenges regarding institutional support emphasized a lack of support from the library, center, or larger parent institution. Concerns included lack of infrastructure to address curation needs, support for data reference for discovery or data literacy, and lack of preservation initiatives. Participants also mentioned a lack of relevant courses currently taught besides government data in MLIS programs, cross-institutional coordination, and solidifying where library responsibilities lie. Further concerns included internal territories, workflows, and demonstrating value to the campus.

Finally, a very insightful concern was that librarians, as individuals, do not follow best practices for managing data. For example, there was a response from a participant—not a librarian—who reflected many data librarians' concerns about demonstrating and communicating value or competency with researchers:

What challenges to data curation do you experience at your institution? "None. The faculty are smart enough to arrange their own data curation and don't really need the library's help."

Another participant clarified what types of institutional support were needed:

"decentralized support for research, lack of clear policies on research data (these are being developed), focus on supporting grants while they are active rather than managing the assets once grants are completed, lack of sufficient funding to support data curation, difficulty communicating across a large and decentralized campus."

Those who felt positively about their success with data curation services attributed it to faculty engagement, enthusiastic environments, institutional support, adequate personnel, and high-quality technology or tools. Other sentiments included that their institution valued their work, that they felt they had a satisfactory master of skill, and that their institution had an exploratory culture. A sense of collaboration frequently associated with institutional support, which could mean a correlation exists between the two. There was also a frequent association between library participation and enthusiasm or prominence (awareness, demand), perhaps showing results from proactive approaches to administration and service. These frequent associations reflect statements in open-ended responses.

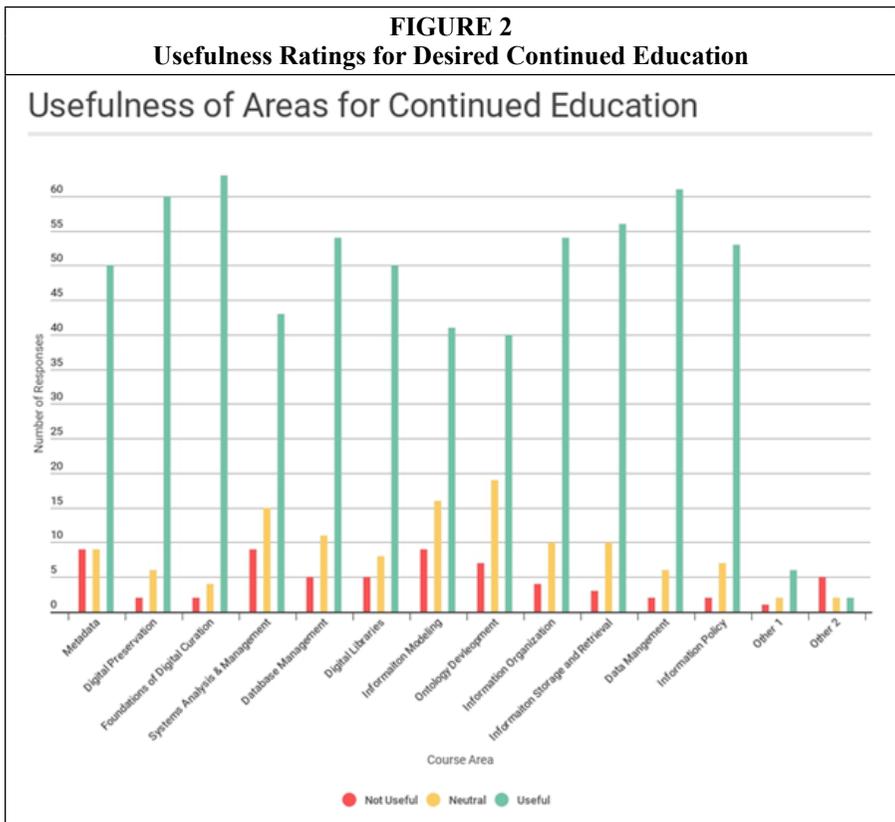
After observing what was currently successful, professionals foretold what might be successful. The most common responses about desired services and future trends in

the next two to five years were cultural changes and federal policies. Others included library participation in data science, self-deposit of data, storage options, and concerns and visualization. Several responses mentioned an increase in professional development and job opportunities, as well as increased scholarship on the specialization. For example:

“As someone with a research, rather than library, background, I had no idea what ‘data curation’ was before I started my current job. This is not something that is integrating into the research workflow. But it is something that we should be thinking about—how much and what kinds of documentation researchers should be collecting during their projects to make sure their data are useable (even to themselves) in the future.”

Professionals noted which knowledge areas in which they desired continuing education. The most desired areas of those listed included Data Management, Information Policy, Digital Preservation, and Information Organization. Additional responses from open-ended questions included Administrative Management, discipline-specific Communication and Scientific Computing. Further areas included: Information Management, Statistics, Programming (coding), Advanced Metadata Standards, Advanced Information Technology and Applied Research Skills (data for disciplinary use, not library use).

FIGURE 2
Usefulness Ratings for Desired Continued Education



When asked if participants were satisfied with the data curation positions or procedures at their institutions, 44.12 percent of participants said Yes and 55.88 percent said No (n = 69). As a follow-up, participants explained that they were laying a solid foundation in a growing area or that there was not enough technological or administrative support. For example, one participant stated: "We are not yet where we need to be. Progress is slow, and there is not (and cannot be) a 'one size fits all' solution."

Additionally, 51.33 percent of participants (n = 71) reported that there was no assessment for data curation services at their institution. Only 18.57 percent said they performed any evaluation, with audits and general user feedback among the methods of assessment. Although it is understandable in a growing area, it is insightful that there are not many initiatives to demonstrate the value of services. Despite the literature and employment outlooks stressing data curation as a crucial need, there were not many instances of service assessment to reflect how professionals are addressing user needs.

Overall, the quantitative and qualitative data from those working in libraries was often neutral or optimistic. Perceptions of those who worked outside libraries were often pessimistic. The investigators believe professionals were reluctant to give a negative opinion, based on qualitative responses. Open-ended responses had as much neutral sentiment as quantitative responses.

In contrast, librarians working with data, including those with backgrounds in other data-driven subjects, believe librarians are currently working well with the resources at hand. Those with outside perspectives did not see libraries as sources of aid with data or that their services were needed. Many working in libraries stressed the need to be able to understand data practices among various disciplines to better communicate their services to researchers and administrators.

Discussion and Conclusion

The results suggest that professionals regard a matriculated MLIS program that includes field experience as the best way to give professionals a sense of qualification for employment in data curation, especially if they do not have other education in a data-driven discipline.

Professionals often perceived their education as out of date. It may be that their training did not meet expectations of directly applicable or future-proof preparation. Perhaps this is why libraries are currently participating in steering committees rather than offering curation services. Additionally, it shows how previous studies that compared data curation course descriptions to data librarian job descriptions did not provide sufficient information. In this study, we see that most data librarians did not attend data curation MLIS programs, yet they were hired as qualified applicants. Instead of looking at aspirational information, we examined what skills were applied in practice. Most of those skills were learned through applied research or other disciplines or were "soft" skills such as communication.

Many did not recognize any indirect or soft skills learned through the MLIS, but some did. Underestimation of soft skills may reinforce the argument that many librarians lack confidence or flexibility, despite interest and continuing education opportunities. We asked participants about educational experiences other than the MLIS, such as certificates or continuing education. Responses from those who took continuing education offerings were very low, suggesting a lack of funding, time, or value. Like MLIS programs, continuing education often centers on discussions more than hands-on experiences.

There may be misplaced blame or self-victimization rather than the proactive attitude among those who felt satisfied with their success so far. Remember, the few professionals who did specialize in MLIS data curation programs did not feel their

education directly prepared them for the workforce either. Several professionals who had negative sentiments about their education or current institutional support were optimistic about future developments. They did not seem to expect direct skills from their education. Expectations of a directly applicable, future-proof and hands-on education may be too rigid, considering the limited resources within MLIS programs. There seems to be a discrepancy among stakeholders (students, instructors, employers, and the like) about expectations for gaining practical experience through education networks or on an individual basis. Certainly, for many, the reality is that individual access to practical experience external to MLIS programs or full-time employment is scarce. In any case, most professionals reported learning more skills on the job (that is to say, hands-on).

Several professionals suggested MLIS education requires hands-on learning with data as a solution. Data professionals often suggested that MLIS programs require all students, whether specializing in data curation or not, to conduct research that included handling and analyzing data. It appears a dedicated specialization is not necessary for each MLIS program to provide. Instead, MLIS programs should focus on offering courses that offer practical skill sets such as research methods and statistical analysis. In the future, data services may evolve to affect all academic librarians. Since many other graduate programs require students to become familiar with applied research, it may behoove LIS programs to do as professionals suggest. It could significantly improve relationships and communication with scholars in the future.

In conclusion, most professionals believe LIS education gave them a solid foundation to then learn on the job. There is an imperative need to expand investment from libraries and LIS programs in data services. Many felt their education or experience with research data in fields other than LIS was essential. Also, they regarded education from matriculated degrees as the main source of acquiring skill sets relevant to data curation. Future developments in education should adapt applied curriculum from data-driven disciplines, for technical as well as relational training. Alternatively, iSchools and LIS programs would do well to build partnerships with data centers and other "data science" degree programs to create "immersive," embedded practical education like the model examined by the University of Pittsburgh with Carnegie Mellon University Libraries and used by the University of Illinois. An immersive approach may also create communication skills for building partnerships with researchers for services. Communication with researchers was the skill that participants desired most. This could demonstrate to students and data professionals in other fields that librarians are qualified to aid them. It could also gain support from administration, which would show there is a great deal of overlap in knowledge areas and an ideal profession to facilitate interdisciplinary data curation support. Moreover, it could strengthen a comprehensive response to researchers, students, funding agencies and the public who need curation to process information better in the time of big data.

Acknowledgements

Special thanks to the Department of Technology and Digital Scholarship at Florida State University Libraries for piloting the survey. Thanks to Ryan and Joni Litsey for local review.

APPENDIX A. Additional Figures and Tables

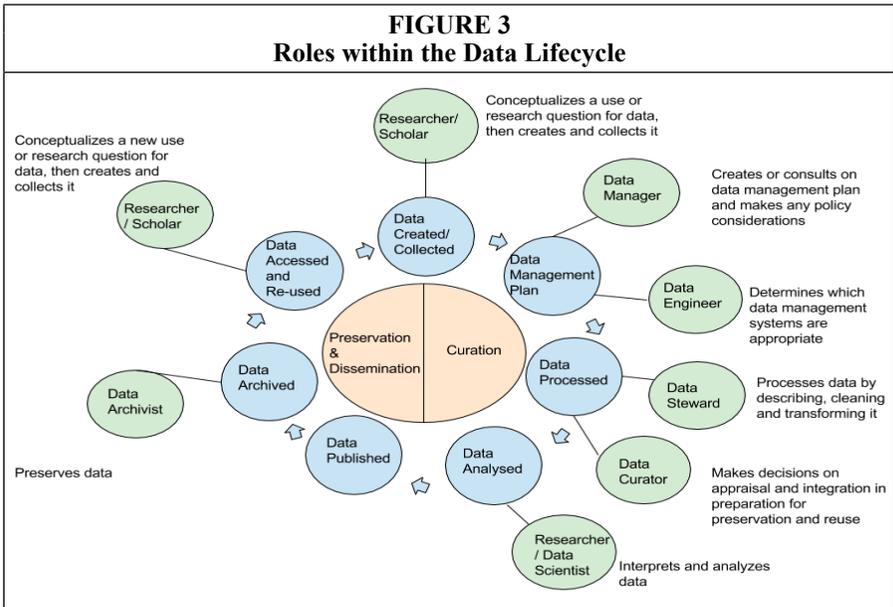


TABLE 6
Other Universities Attended for MLIS or Relevant Programs

Other Universities	Number of Responses (n = 63)
Indiana University (US)	6
University of North Texas (US)	4
Dalhousie University (Canada)	3
University of Wisconsin–Madison (US)	2
State University of New York–Buffalo (US)	2
Pratt Institute (US)	2
Sapienza University–Rome (Italy, EU)	2
University of California–Los Angeles (US)	2
Kent State University (US)	2
Oregon State University (US)	2
Florida State University (US)	2
North America: University of William and Mary, Case Western University, University of Minnesota, Stevens Institute of Technology, San Jose State University, University of Wisconsin–Milwaukee, University of North Carolina–Greensboro, Louisiana State University, Valdosta State University, Yale, New York University, University of Pittsburg, St. John’s University, University of California–Santa Cruz, California Polytechnic, University of Washington, City University of New York–Queens, Catholic University of America, University of Missouri–Columbia, McGill University (Canada), Syracuse University, University of Wisconsin–Oshkosh	>2

TABLE 6	
Other Universities Attended for MLIS or Relevant Programs	
Other Universities	Number of Responses (n = 63)
United Kingdom and Europe: Delft University of Technology (UK), University of Essex, Cambridge University, University of the West of England, University of Leiden (Netherlands), Wageningen University and Research Centre (Netherlands), University of Zurich (Switzerland), Swiss Federal Institute of Technology in Zurich/ ETH Zurich	>2
Australia and New Zealand: Monash University	>2

APPENDIX B. Survey Questions

1. How many full-time employees (FTE) are dedicated to data curation at your institution?
2. What data curation responsibilities are part of the services you provide? Check all that apply.
 - a. creating or managing data systems
 - b. preserving data
 - c. using or modifying web 2.0 tools
 - d. developing workflows
 - e. consulting on data management plans and compliance for researchers
 - f. analyzing or interpreting data
 - g. developing applications
 - h. documenting procedures
 - i. developing policies
 - j. ensuring security or quality control
 - k. metadata standards and/or manipulation of metadata
 - l. programming (languages)
 - m. promoting open access or data sharing
 - n. teaching data literacy and best practices
 - o. "data interview" or needs assessment
 - p. consulting on intellectual property or copyright issues
 - q. providing access support
 - r. managing or training personnel
 - s. other
3. Check all the types of data you or your institution works with:
 - text images
 - presentations
 - videos
 - spreadsheets
 - databases
 - spatial data
 - computational models
 - computational code spreadsheets
 - other
4. Which type of information institution employs your data curation professionals?
 - a. Public Library
 - b. Academic Library

- c. College/Community College Library
 - d. Private Company
 - e. Research Center
 - f. Government
 - g. Nonprofit organization
 - h. Other
5. What levels of education do those employees have?
 - a. High School/GED
 - b. Associate Degree
 - c. Bachelor's Degree
 - d. Master's degree
 - e. Specialists Degree
 - f. PhD
 - g. Other
 6. Which university did the professional attend?
 - a. University of Illinois–Champaign-Urbana
 - b. University of Maryland
 - c. Rutgers University
 - d. University of North Carolina–Chapel Hill
 - e. University of Arizona
 - f. University of Michigan
 - g. Simmons College
 - h. Other
 7. How well do graduate library and information programs prepare graduates for their jobs?
 - a. very well
 - b. well
 - c. neutral
 - d. fair
 - e. poorly
 8. How has the education equipped the professional(s) for curating data at your institution?
 9. Was their education specialization specifically in data curation?
 - a. yes
 - b. no
 10. What type/level of certification do those employees have if any?
 - a. specialization within a matriculated degree
 - b. graduate certificate
 - c. workshop certification
 - d. mass open course certification
 - e. professional development or continued education certification
 - f. other
 11. Select any of the following courses did the professional took.
 - Metadata in Theory
 - Digital Preservation
 - Foundations of Data Curation
 - Information Processing
 - Systems Analysis and Management
 - Database Management
 - Digital Libraries
 - Information Modeling

- Ontology Development
 - Information Organization
 - Information Storage and Retrieval Other
 - Other
12. Rate courses from most to least useful (scale of 1 to 5, 1 being least useful and 5 being most useful)
- Metadata in Theory
 - Digital Preservation
 - Foundations of Data Curation
 - Information Processing
 - Systems Analysis and Management
 - Database Management
 - Digital Libraries
 - Information Modeling
 - Ontology Development
 - Representing and Organizing Information Resources
 - Information Storage and Retrieval
13. What types of courses do you think should be added to data curation program curriculums?
14. What challenges exist at your institution concerning data curation practices?
15. What aspects of your institution aid in your success with data curation practices?
16. Select which areas of continued education would be beneficial for the position?
- a. policy development
 - b. data interpretation
 - c. consulting/liasing
 - d. preservation systems
 - e. selection and appraisal
 - f. formats, standards, and practices
 - g. modeling
 - h. training (management)
 - i. other
17. Are you satisfied with the data curation position(s) or procedures available at your institution?
- a. yes
 - b. no
18. What developments or trends have you observed in the past 2–5 years in data curation?
19. What other data curation services would you like to perform for patrons?
20. How does your institution evaluate data curation services?

Notes

1. Rebecca L. Harris-Pierce and Yan Quan Liu, "Is Data Curation Education at Library and Information Science Schools in North America Adequate?" *New Library World* 113, no. 11/12 (2012): 598–613.
2. P. Bryan Heidorn, "The Emerging Role of Libraries in Data Curation and E-science," *Journal of Library Administration* 51, no. 7/8 (2011): 662–72.
3. Committee on Future Career Opportunities and Educational Requirements for Digital Curation, Board on Research Data and Information, Policy and Global Affairs and National Research Council, *Preparing the Workforce for Digital Curation* (The National Academies Press: 2015), 1–79.
4. Interuniversity Consortium for Political and Social Research (ICPSR), "Data Management & Curation," available online at <https://www.icpsr.umich.edu/icpsrweb/content/datamanagement/> [accessed 12 January 2017].

5. Elaine Martin, "What Is Data Literacy?" *Journal of eScience Librarianship* 3, no. 1 (2014): 1–2. doi:10.7191/jeslib.2014.1069.
6. Ibid.
7. ICPSR, "Data Management & Curation."
8. *Preparing the Workforce for Digital Curation*, 10–11.
9. Cynthia Hudson-Vitale et al., *SPEC Kit 354, Data Curation* (Chicago, Ill.: Association of Research Libraries, 2017), 141.
10. Bill Howe, "Big Data Curricula at the UW eScience Institute" (presented at the Joint Statistical Meeting, August 3–8, 2013, Montréal Québec, Canada).
11. Digital Preservation Coalition, "Glossary," in *Digital Preservation Handbook* (2008), available online at <http://handbook.dpconline.org/introduction> [accessed 10 September 2016].
12. Margaret Hedstrom, "Digital Preservation: A Time Bomb for Digital Libraries," *Computers and the Humanities* 31, no. 3 (1997): 189–202.
13. Heidorn, "The Emerging Role of Libraries in Data Curation and E-Science."
14. *Preparing the Workforce for Digital Curation*, 10–11.
15. Joyce L. Ogburn, "The Imperative for Data Curation," *Libraries and the Academy* 10, no. 2 (2010): 241–46; Megan Toups and Michael Hughes, "When Data Curation Isn't: A Redefinition for Liberal Arts Universities," *Journal of Library Administration* (2013): 223–33.
16. *SPEC Kit 354, Data Curation*, 10–21.
17. Anna Gold, "Data Curation and Libraries: Short-Term Developments, Long-Term Prospects" (paper presented at American Geophysical Union annual meeting, December 14–18, 2009, San Francisco, CA), available online at http://digitalcommons.calpoly.edu/lib_dean/27/ [accessed 7 March 2018].
18. Si Li, Zhuang Xiaozhe, Xing Wenming, and Guo Weining, "The Cultivation of Scientific Data Specialists: Development of LIS Education Oriented to E-Science Service Requirements," *Library Hi Tech* 31, no. 4 (2013): 700–24.
19. *SPEC Kit 354, Data Curation*, 10–21.
20. Stephanie L. Maatta, "The Emerging Databrarian: Digital and Data Management Roles Reach Critical Mass among Recent LIS Graduates," *Library Journal* 138, no. 17 (2013): 26–33.
21. Suzie Allard, "Placements & Salaries 2016: Bouncing Back," *Library Journal* (Oct. 17, 2016), available online at <https://lj.libraryjournal.com/2016/10/placements-and-salaries/2016-survey/bouncing-back/> [accessed 7 March 2018]; Suzie Allard, "Placements & Salaries: 2016 Explore All the Data," *Library Journal* (Oct. 17, 2016), available online at <http://lj.libraryjournal.com/2016/10/placements-and-salaries/2016-survey/explore-all-the-data-2016/> [accessed 7 March 2018].
22. Ogburn, "The Imperative for Data Curation," 241–46.
23. *Preparing the Workforce for Digital Curation*, 53–57.
24. Si Li et al., "The Cultivation of Scientific Data Specialists."
25. Carole L. Palmer, Cheryl A. Thompson, Matthew S. Mayernik, Valerie Williams, Karon Kelly, and Suzie Allard, "Data Curation Education in Research Centers: Formative Evaluation Findings from Years One and Two." Poster presented at the 9th International Digital Curation Conference, February 24–27, 2014, San Francisco, CA. Available online at <http://hdl.handle.net/2142/48950> [accessed 14 February 2018].
26. Carol Tenopir, Robert J. Sandusky, Suzie Allard, and Ben Birch, "Research Data Management Services in Academic Research Libraries and Perceptions of Librarians," *Library & Information Science Research* 36, no. 2 (2014): 84–90.
27. *SPEC Kit 354, Data Curation*, 10–21.
28. *Preparing the Workforce for Digital Curation*, 17.
29. Tenopir et al., "Research Data Management Services," 88–90.
30. Benjamin B. Branch, Kristin Wegner, Steven Smith, Darrell Schulze, Venkatesh Merwade, and Jinha Jung, "A Library Approach to Establish an Educational Data Curation Framework (Edcf) That Supports K–12 Data Science Sustainability." Poster presented at the American Geophysical Union Fall Meeting, December 9–13, 2013, San Francisco, California. Available online at https://docs.lib.purdue.edu/lib_fspress/64/ [accessed 14 February 2018].
31. Ogburn, "The Imperative for Data Curation," 241–46.
32. Palmer et al., "Data Curation Education in Research Centers"; Ogburn, "The Imperative for Data Curation."
33. Nicholas M. Weber, Carole Palmer, and Tiffany C. Choa, "Current Trends and Future Directions in Data Curation Research and Education," *Journal of Web Librarianship* 6, no. 4 (2012): 305–20.
34. Ibid.
35. *Preparing the Workforce for Digital Curation*, 53–57.
36. Dorothea Salo, "Competency Lists Considered Harmful: Can We Rethink Them? Peer-to-Peer Review," *Library Journal* (2014), available online at <http://lj.libraryjournal.com/2014/05/>

opinion/peer-to-peer-review/competency-lists-considered-harmful-can-we-rethink-them-peer-to-peer-review/ [accessed 19 March 2018].

37. Howe, "Big Data Curricula at the UW eScience Institute."
38. Koltay Tibor, "Data Literacy: In Search of a Name and Identity," *Journal of Documentation* 71, no. 2 (2015): 401–15.
39. *Preparing the Workforce for Digital Curation*, 53–57.
40. Allard, "Placements & Salaries."
41. *Preparing the Workforce for Digital Curation*, 54.
42. Tenopir et al., "Research Data Management Services," 86; Jake Carlson, "Opportunities and Barriers for Librarians in Exploring Data: Observations from the Data Curation Profile Workshops," *Journal of eScience Librarianship* 2, no. 2 (June 20, 2013): 17–33, doi:10.7191/jeslib.2013.1042.
43. Weber, Palmer, and Choa, "Current Trends and Future Directions," 305–20; Tibor, "Data Literacy," 401–05; *Preparing the Workforce for Digital Curation*, 53–57.
44. Mandi Goodsett and Amanda Koziura, "Are Library Science Programs Preparing New Librarians? Creating a Sustainable and Vibrant Librarian Community," *Journal of Library Administration* 56, no. 6 (2016): 697–721.
45. Carole L. Palmer, "Priorities for Data Curation Education: Data Center Partnerships and Long-Tail Science," in *Microsoft eScience Workshop Collection* (2012) held October 8–9, 2012, at the Hyatt Regency Chicago in conjunction with the IEEE International Conference on eScience 2012.
46. Virgil E. Varvel, Jr., Elin J. Bammerlin, and Carole L. Palmer, "Education for Data Professionals: A Study of Current Courses and Programs." Paper presented at the 2012 iConference Toronto, Ontario, Canada, February 7–10, 2012 .
47. Ibid.
48. Goodsett and Koziura, "Are Library Science Programs Preparing New Librarians?" 712–14.
49. Si Li et al., "The Cultivation of Scientific Data Specialists"; Harris-Pierce and Liu, "Is Data Curation Education at Library and Information Science Schools in North America Adequate?"
50. Varvel et al., "Education for Data Professionals."
51. *Preparing the Workforce for Digital Curation*; Palmer et al., "Data Curation Education in Research Centers."
52. Cheryl A. Thompson, Megan Senseney, Karen S. Baker, Virgil E. Varvel, and Carole L. Palmer, "Specialization in Data Curation: Preliminary Results from an Alumni Survey, 2008–2012," Proceedings of the Association for Information Science and Technology, May 8, 2014, available online at <http://onlinelibrary.wiley.com/doi/10.1002/meet.14505001151/full> [accessed 5 June 2015].
53. Ibid.
54. Goodsett and Koziura, "Are Library Science Programs Preparing New Librarians?" 706–11.
55. Lili Luo, "Fusing Research into Practice: The Role of Research Methods Education," *Library & Information Science Research* 33, no. 3 (2011): 191–201.
56. Doug Henschen, "Big Data Analytics Master's Degrees: 20 Top Programs," *InformationWeek* (January 7, 2013). Available online at <https://www.informationweek.com/big-data/big-data-analytics/big-data-analytics-masters-degrees-20-top-programs/d/d-id/1108042> [accessed 14 February 2018].
57. Il-Yeol Song and Yongjun Zhu, "Big Data and Data Science: What Should We Teach?" *Expert Systems* 33, no. 4 (2015): 364–73.
58. Ibid.
59. Yuri Demchenko and Wouter Los, "Edison: Education for Data Intensive Science to Open New Science Frontiers," *Mena Report* (2015): 1–5.
60. Howe, "Big Data Curricula at the UW eScience Institute."
61. Song and Zhu, "Big Data and Data Science."
62. Charles Pensig, "Why I Left My Data Science Master's Program." LinkedIn Pulse, May 19, 2015. Available online at <https://www.linkedin.com/pulse/why-i-left-my-masters-program-charles-pensig-1> [accessed 14 February 2018].
63. Toupes and Hughes, "When Data Curation Isn't," 223–33.
64. Demchenko and Los, "Edison: Education for Data Intensive Science," 3–4; Thompson et al., "Specialization in Data Curation."
65. Henschen, "Big Data Analytics Master's Degrees: 20 Top Programs;" "Directory of ALA-Accredited and Candidate Programs," available online at <http://www.ala.org/educationcareers/accreditedprograms/directory> [accessed 14 February 2018]; Song and Zhu, "Big Data and Data Science."