Issues in Science and Technology Librarianship Summer 2017
DOI:10.5062/F4V40SFK

Refereed article

Comparison of Data and Informatics Responsibilities and Job Titles between Academic STEM and Medical Librarians

Robyn B. Reed
Assistant Librarian, Biomedical Informatics and Emerging Technologies
Harrell Health Sciences Library
Penn State College of Medicine
Milton S. Hershey Medical Center
Hershey, Pennsylvania
rreed4@pennstatehealth.psu.edu
http://orcid.org/0000-0003-2141-0490

Nancy J. Butkovich
Associate Librarian, Head, Physical and Mathematical Sciences Library
The Pennsylvania State University
University Park, Pennsylvania
njb2@psu.edu
http://orcid.org/0000-0002-9188-6518

Abstract

Discussions abound regarding current and future roles of academic science and medical librarians. As changes in scientific approaches, technology, scholarly communication, and funding mechanisms occur, libraries supporting scientific areas must be equipped to handle the various needs of these researchers. The purpose of this study was to examine how academic STEM (Science, Technology, Engineering, Mathematics) and medical libraries are responding to these changes, specifically in the areas of data and informatics support. An examination of job descriptions between 2005 and 2014 showed an increase over time in the number of advertisements with responsibilities in data or informatics. Our study suggests STEM and medical librarians are responsible for many different data-intensive areas, mostly in data management or curation, although librarians filling many of these positions also had traditional librarianship responsibilities, such as general reference or collection development. The hybrid nature of these positions was also reflected in the job titles, of which nearly half did not indicate that the position had significant data or informatics-related responsibilities.

Introduction

The roles academic librarians play in supporting scientific research are highly varied and expanding. Technological advances and the increase in data-driven science contribute to evolving scientific methodologies. Science and medical librarians are responsible for many different activities in their jobs. Some of these include traditional library duties, while other new or emerging duties include working in data-intensive areas, such as the need to store, manage, and curate data in multiple forms. In 2013, Cooper and Crum found that between 1990 and 2012 health sciences libraries witnessed the development of new roles as well as observing changes in existing functions such as instruction, outreach, and clinical librarian roles. New positions included bioinformationists, emerging technologies, and data management librarians (Cooper & Crum 2013).

Institutional and library support for data-driven science is growing. Anderson et al. reported that academic biomedical researchers indicated they needed support with data management as well as conducting analyses of their large-scale data (Anderson et al. 2007). A few areas are well established. Gold characterized librarian data roles in Geographic Information Systems (GIS), social science, and bioinformatics and suggested librarians expand into data curation, data consultancy, data publishing, and data reuse (Gold 2007). Two studies by Tenopir et al. suggested that, while some academic libraries are offering research data services, many are still in the planning stages (Tenopir et al. 2015; Tenopir et al. 2014). Antell et al. found that some academic librarians have roles in research data support including data or institutional repositories (Antell et al. 2014).

The terms "e-science" and "e-research" reflect efforts to address data and information needs and are being used by libraries and research institutions alike. A report on e-science and data services among Association of Research Libraries (ARL) members described e-science as including computational science, team science, and networked science across all scientific domains. The study indicated that 73% of the ARL libraries surveyed were involved in e-science at their institutions while others were interested in supporting e-science and were making plans to do so (Soehner et al. 2010).

STEM and medical librarians share overlapping challenges with research data, yet they serve different patron populations. The goals of this study were to determine data and informatics-related aspects of these librarians' jobs and to identify differences in these roles between these populations of professionals. We were also interested in determining if the job titles in the advertisements in our data set reflected changes in the job responsibilities of STEM and medical librarians.

Methods

The use of job advertisements for research studies to ascertain trends in the field is common in librarianship. While many studies focus on training and experience needed for library positions, others examine job responsibilities to understand changing roles of librarians (Antell et al. 2014; Han & Hswe 2010; Meier 2010). We obtained job descriptions by searching five publicly available mailing lists containing medical or STEM librarian advertisements: MEDLIB-L (Medical Libraries Discussion List, Medical Library Association), STS-L (Science and Technology Section ACRL/ALA), PAMNET (Physics Astronomy Mathematics Division SLA), CHMINF-L (Chemical Information), and ELDNET-L (Engineering Libraries Division ASEE). We reviewed all posts for job advertisements from 2005-2014 and identified relevant position descriptions to include in our final data set.

We restricted the study to permanent, full-time academic positions in the United States, and excluded positions that were primarily administrative, internships, or temporary assignments, non-medical or non-STEM jobs. Also excluded were advertisements where data or informatics responsibilities comprised less than 25% of the duties. This assessment was based on how responsibilities were described and emphasized in the announcements. We divided the job descriptions into STEM or medical categories based on employment location. Medical positions were associated with medical schools, while STEM positions were either in academic general or science libraries.

In reviewing advertisements, we looked for terms or activities associated with data and informatics. Examples of these included various data-related terminology (e.g., data management, data mining), bioinformatics, cheminformatics, working with data repositories, digital projects, or initiatives, and terms associated with organizing knowledge such as controlled vocabularies, ontologies, or metadata. A complete listing is in Appendix A.

Results

We examined an estimated 1,100-1,200 advertisements. The majority either did not mention any data or informatics responsibilities, and of those that did, many indicated those duties were a minor part of the job. The precise number was impossible to determine due to duplicate posts on different mailing lists, reminders to apply for positions already posted, advertisements that were withdrawn and revised before being reposted, were neither STEM nor medical positions, or were not in academic institutions. The final data set consisted of 100 job advertisements with data or informatics duties. Figure 1 illustrates the distribution of 70 STEM and 30 medical positions from 2005-2014.

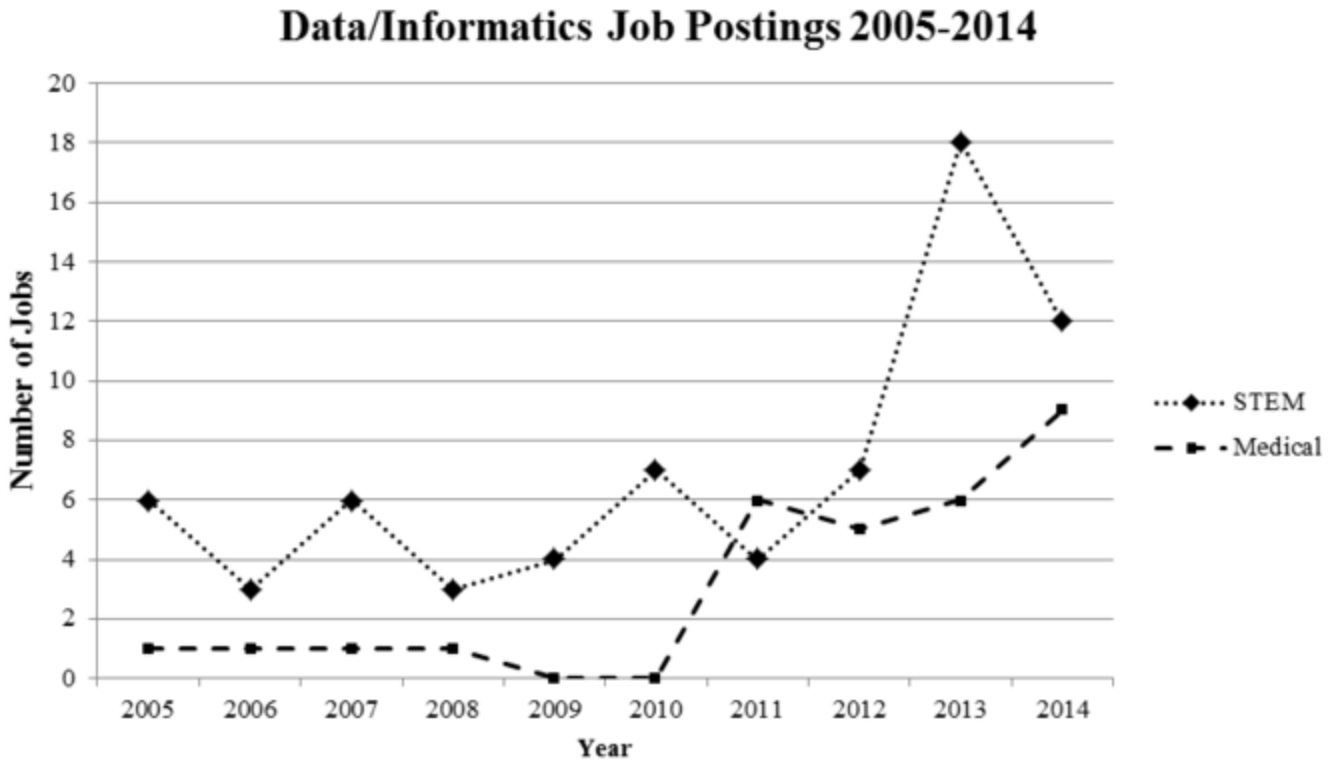

Figure 1. Academic library job postings with significant data or informatics duties

Although the numbers of positions fluctuate from year to year, the overall range for STEM and medical positions remained relatively stable until 2011-2012, when the numbers began rising well above the ranges for the earlier years, although STEM numbers dropped in 2014. A future study will be necessary to determine if the trend toward higher numbers continues. In our study, we expected to find more STEM jobs compared to medical because there are many more academic positions that support STEM areas than there are medical library jobs. Due to the low number of qualifying ads in the early years of our study, we did not analyze the results of the study by year but rather looked at the decade as a whole.

A goal of this study was to understand the types of data-related duties librarians are performing. A typical job description may include functions across a variety of traditional responsibilities such as reference, instruction, collection development, liaison, and other capacities as well as data-related duties. Table 1 illustrates details of the responsibilities within the job descriptions in our data set. Tasks were placed either into broad traditional job duty areas, such as reference or instruction, or into a data or informatics-related duty area, such as data services/digital scholarship or scholarly communication.

Table 1. Percentage of job advertisements by duty type

Job Duty Area	Data/Informatics Duties by Percentage (n=70)		General Librarianship Duties by Percentage (n=30)	
	STEM	Medical	STEM	Medical
Data services/Digital scholarship	94	73	3	7
Instruction	34	63	64	53
Scholarly communication	34	37	1	0
Reference/consultation	33	50	64	63
Liaison/subject specialist	27	37	63	53
Collection development	17	10	57	33
Sponsored research	14	33	6	17
Management/administrative/supervisory	7	7	13	7
Cataloging	4	3	4	7

The values in Table 1 represent the percentage of jobs with roles in the indicated job duty areas. Over the entire study period, data services, instruction, reference, and liaison roles occurred most frequently in our data set. Job duty areas in each advertisement could be assigned to either or both data/informatics and general categories, potentially resulting in percentages exceeding 100%. For example, within the reference/consultation job duty area, an announcement might specify that the librarian could work regular reference desk hours (general) as well as consult with researchers about data management (data/informatics). Not included in the analysis were requirements of trend awareness in various areas, committee work, and professional development.

Considering the criteria used to select the advertisements for this study, finding heavy librarian involvement in various data services as well as data-related instruction, reference, scholarly communications and liaison areas was not surprising. We observed less emphasis on duties in data-related collection development, cataloging, sponsored research, or supervisory roles. Our results also showed many of these librarians perform library duties for traditional audiences, such as general reference or library instruction for undergraduates, in addition to having specialized responsibilities in data or informatics.

In order to determine the balance between traditional and data-specific duties, we assigned postings to one of two percentage ranges based on how data-intensive an advertisement appeared to be relative to the total responsibilities and emphasis given to the data/informatics duties. Table 2 notes the number of advertisements that are more (≥50%) or less (25-50%) data-intensive. Medical positions had a slight edge over STEM with regard to the percentage of ads with ≥50% data/informatics duties.

Table 2. Data/informatics duties by percentage range

	Number (Percent) of STEM Positions (n=70)	Number (Percent) of Medical Positions (n=30)
Ads with Data/Informatics Duties ≥ 50%	32 (46%)	16 (53%)
Ads with Data/Informatics Duties 25-50%	38 (54%)	14 (47%)

Because data services comprised large percentages of the job descriptions in our data set, we took a closer look at the types of services described in this area. They are shown in Figure 2. Most job listings included general terms such as "data services" but many provided detailed information as well. Redundancy in the figure totals occurred as most job descriptions incorporated several duties. The most common subject areas identified in the data are noted in Figure 2.

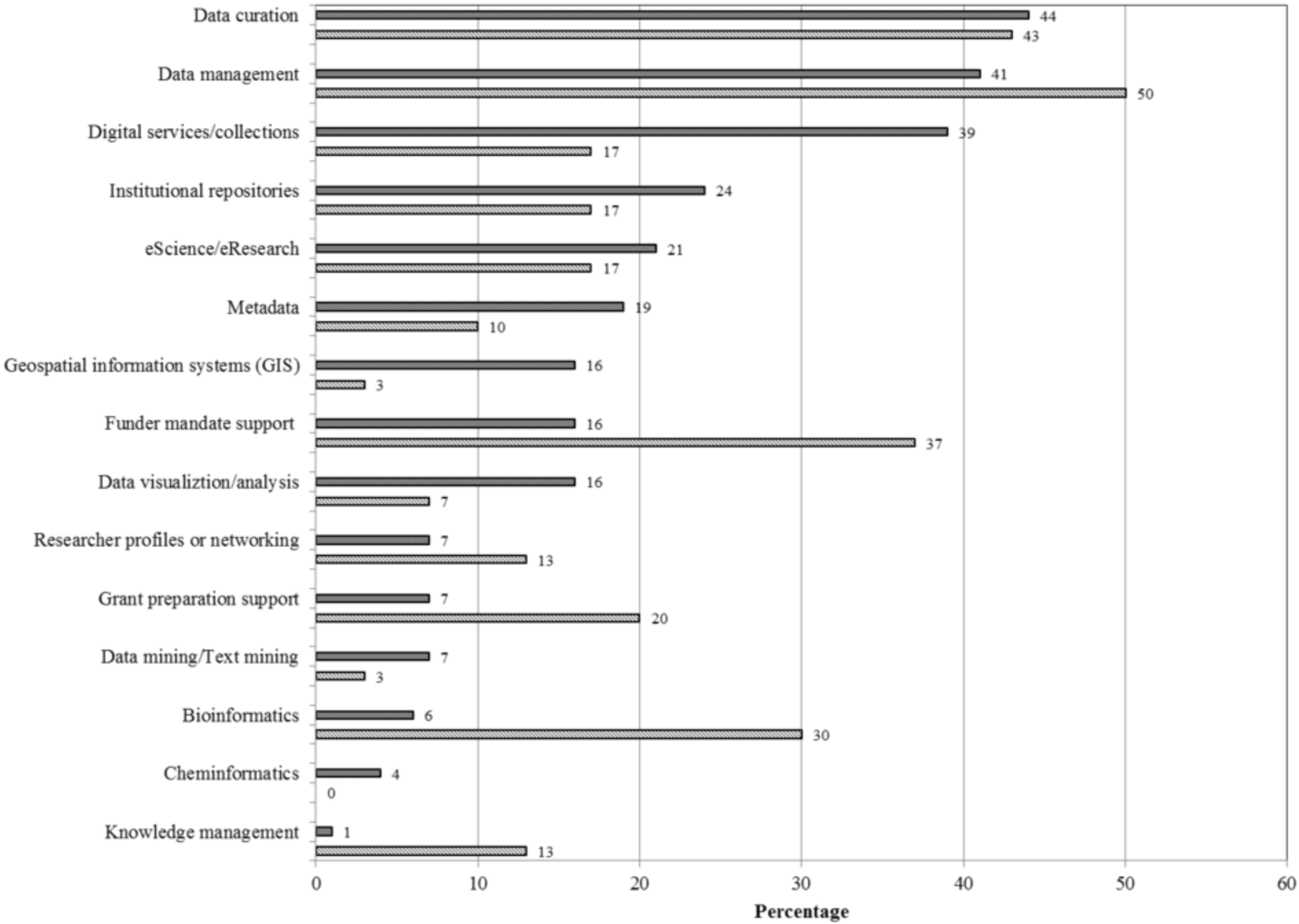

Figure 2. Data services by category and percentage of jobs; STEM (n=70; dark gray bar), and Medical (n=30; light gray bar)

Some vague terminologies in data services were difficult to sort into precise subcategories, as typical descriptions included phrases that indicated candidates would be responsible for areas including collection and preservation of data without including much detail. Duties that were challenging to place were counted in either the digital services and collections or data management categories, depending on context. The digital services and collections grouping comprised responsibilities working in areas including digital projects, digital libraries, and digital discovery. Data management is a term often used independently, yet in this analysis may have involved some data discovery and access as well. Institutional repository (IR) duties included the development of the IR system in addition to working with the data, metadata, or collections in the IR. The metadata category generally consisted of creating metadata for data sets. Work with researcher profiles sometimes named the platform, such as VIVO, being used by the institution for showcasing research output. When present, the term "e-science" or "e-research" appeared without specific definitions. Alvaro et al. (2011) reviewed areas that encompass e-science.

While demonstrating that some academic librarian positions are more data- or informatics-intensive, we questioned if the corresponding job titles reflected this. Despite their specialized characteristics, the number of positions described as "specialists" was low. Over the ten-year period examined, "specialist" was only used on 14% of the jobs while "librarian" was used 76% of the time. Moreover, some positions employed both terms, e.g. librarian/specialist. This resulted in some redundancy in the observed percentages. The few remaining positions included titles of informationist, informaticist, consultant, associate, or head. In addition, the job title analysis revealed that 54% of the titles included a data/informatics term while 46% of them did not. Of note was the similarity between STEM and medical position descriptions, where 51% of STEM library and 60% of medical library titles in the data set contained a data or informatics term. The observation was very similar when examining general titles, where 49% of STEM and 40% of medical position titles were general in nature.

Some of the unique or noteworthy titles included: e-science librarians (twice in STEM), big data information specialist (once in STEM), bioinformaticist (once in medical), informationists (twice in medical, once in STEM), translational science/research librarian/specialist and bioinformatics librarian/specialist (four times each in medical). Titles with variant forms, such as data services, research services, GIS, spatial data, research data, science data, and data management, occurred regularly in librarian and specialist titles.

Discussion

Our goal was to examine similarities and differences between medical and STEM librarian positions with data and informatics responsibilities. Both showed an increase in data/informatics positions in recent years (Figure 1), yet many of the positions contained general library duties as well as specialized data/informatics responsibilities (Table 1). Tenopir et al. (2015) examined research data services growth in academic libraries and found little change in the services offered. Our results were mostly consistent with these conclusions, as medical and STEM libraries job descriptions indicate support for these areas but do not always provide details about how this will occur.

Research funding requirements of the National Science Foundation (NSF) and the National Institutes of Health (NIH) may have contributed to the increase in data- and informatics-related jobs in the timeframe examined. In 2008 the NIH issued a Public Access Policy requiring NIH-funded, peer-reviewed research to be made publicly available (NIH 2014); policy compliance for continuation grant awards began in 2013 (NIH 2013). In 2011, the NSF began requiring all grant proposals to contain a Data Management Plan (DMP) for disseminating and sharing research data (NSF n.d.). Finally, a 2013 memorandum from the Office of Science and Technology Policy required expanded public access requirements to other large Federal

agencies (Holdren 2013). Our data suggest that libraries are supporting these requirements through assistance with data management, writing DMPs, assisting researchers with public access compliance, and other requirements.

STEM and medical positions differed in the types of data services provided (Figure 2). We frequently encountered data management roles in our data set across both medical and STEM positions (Figure 2). While it is clear libraries are engaging in data management, the exact roles of librarians in this area remain unclear (Si et al. 2013; Antell et al. 2014). GIS, data visualization, and data analysis services were more common among STEM libraries than medical. However, data visualization and data analysis were not always clearly defined within the advertisements, possibly because of the relative newness of these technologies. Bioinformatics roles were more frequent among medical than STEM libraries. Although bioinformatics has applications in STEM, other informatics specializations, such as cheminformatics, may be more relevant.

Our study found libraries are also looking for data curation, e-science, or e-research skillsets. This is consistent with Heidorn (2011), who noted emerging data curation and e-science roles within libraries. He argued that librarians are well qualified to manage research data following minor retraining in digital documents. Considering many ARL libraries offer or are planning to offer e-science services (Soehner et al. 2010), it was logical that we would observe these in our data set. Since the definition of e-science varies, however, it is difficult to ascertain what these duties entail (Alvaro et al. 2011; Gold 2007; Soehner et al. 2010).

Another data services role we identified in our data set was working with metadata. The number of metadata librarian positions has increased in academic settings in recent years (Han & Hswe 2010), however, not all advertised metadata librarian positions met our criteria. We did not include announcements focusing on traditional cataloging metadata; instead, we considered positions dealing with metadata creation and utilization for STEM or medical research data. Additionally, some metadata job postings may not have been counted because they were included within a term like e-science rather than as a separate item.

Other areas of responsibility, such as data or text mining, did not appear frequently in our data set, although they could be included in general terms, such as data support services. Some positions encouraged candidates to be aware of data mining techniques or preferred experience with data mining, so this appears to be an expanding service area.

Multitasking is a growing trend in science librarianship (Meier 2010), and we observed evidence of multitasking in the advertisements in our study. Most postings contained a large number of responsibilities, and in about half of the advertisements, fewer than 50% of the duties were in data/informatics areas (Table 1, 2). This tendency toward multitasking was also reflected in the job titles where 46% of them contained general or broad descriptions of the position, e.g., science librarian. Only 14% of the job titles in our data set described the positions as specialist, and nearly all of these had over half of their job duties in data/informatics areas. These characteristics suggest these positions were traditional librarian jobs with data/informatics responsibilities added to them.

The most frequently appearing job title variants in our study were similar to those identified by Si et al. (2013), with the exception of GIS or spatial data positions, which did not appear on Si's list. Also, our data set showed that medical libraries posted four translational science/research positions from 2011-2012, likely in response to the NIH Clinical and Translational Science Awards, which support biomedical research to improve health outcomes. Currently about 60 U.S. academic institutions have established programs in this area (NIH 2016).

Analyzing job postings is difficult due to inconsistencies in describing the responsibilities, since many terms are not well defined, and vague terms like "data support services" are common. Although we did not examine required or preferred skills, we observed some employers asking for proficiencies that did not appear to connect to the job responsibilities. These factors may indicate employer indecisiveness or reliance on the successful candidate developing specific data services based on his or her skillsets. We are not the first to note the challenges of interpreting ambiguous job advertisements. Hernon and Schwartz (2015) provided an example of how it is difficult to understand what employers are seeking in job postings.

Uncertainty regarding how much post-hiring modifications the actual duties would undergo is one of the limitations of this study. Although the mailing lists we used contained advertisements from all areas of STEM and medical librarianship, there are many other places job announcements can appear, so it is possible some relevant postings were not captured. Finally, we do not know what impact, if any, the economic recession from 2008-2010 had on the number of advertised positions. Regazzi (2013) reported that library staffing during the recession was maintained or expanded in some cases, yet our anecdotal impression was that the number of advertisements dropped, particularly in 2009, the worst year of the recession.

Conclusion

Although there are many similarities in duties between STEM and medical librarians with data or informatics responsibilities, there are some differences as well. For example, medical librarian postings were more likely to have over half of the duties related to data or informatics. Although they performed the same general categories of duties, STEM librarians were more likely to be providing those duties in a general setting rather than one focused on data or informatics. With regard to data or informatics duties, STEM librarians were more likely to be involved with digital services/digital collections, metadata, institutional repositories, data visualization/analysis, and GIS, while medical librarians were more involved in data management, funder mandate support, grant proposal support, bioinformatics, and knowledge management.

Data and informatics duties are being built into existing jobs for STEM and medical librarians as additional responsibilities. At the same time, new jobs focusing heavily on data and informatics are being created. Si et al. (2013) and Antell et al. (2013) both noted the lack of clarity about whether the traditional librarian with data or informatics-related duties or the data/informatics specialist would eventually become the model for these positions. After having concluded this study, however, we believe the answer to this question is not which of these two models will prevail, but rather that both models will remain in use, and the needs of the specific library will determine whether a dedicated data/informatics specialist or a more traditional librarian with some data/informatics duties is needed.

Acknowledgements

The authors thank Esther Dell, MLS, who assisted with NIH Public Access Policy information.

Penn State University's Institutional Review Board (IRB) determined that this study did not meet the definition of human subject research and that IRB review and approval was not necessary (STUDY00001669).

Part of the 2014 job advertisement data were presented at the Medical Library Association (MLA) annual conference in Austin, TX, in 2015. After considerable revisions to search terms and method, some of these results were presented at the MLA Tri-Chapter meeting in Philadelphia, PA, in 2016.

References

Alvaro E., Brooks H., Ham M., Poegel S. & Rosencrans S. 2011. E-science librarianship: field undefined. *Issues in Science and Technology Librarianship*. 66. DOI: 10.5062/F46Q1V55

Anderson N.R., Lee E.S., Brockenbrough J.S., Minie M.E., Fuller S., Brinkley J. & Tarczy-Hornoch P. 2007. Issues in biomedical research data management and analysis: needs and barriers. *Journal of the American Medical Informatics Association* 14(4):478-488. DOI: 10.1197/jamia.M2114

Antell K., Foote J.B., Turner J. & Shults B. 2014. Dealing with data: science librarians' participation in data management at Association of Research Libraries institutions. *College & Research Libraries* 75(4):557-574. DOI: 10.5860/crl.75.4.557

Cooper, I.D. & Crum J.A. 2013. New activities and changing roles of health sciences librarians: a systematic review, 1990-2012. *Journal of the Medical Library Association* 101(4):268-277. DOI: 10.3163/1536-5050.101.4.008

Gold, A.K. 2007. Cyberinfrastructure, data, and libraries, part 2: Libraries and the data challenge: Roles and actions for libraries. *D-Lib Magazine* [Internet]. [Cited 2017 Feb 10]; 13(9/10). Available from: http://www.dlib.org/dlib/september07/gold/09gold-pt2.html

Han M.J. & Hswe P. 2010. The Evolving role of the metadata librarian: competencies found in job descriptions. *Library Resources & Technical Services* [Internet]. [Cited 2017 Feb 10]; 54(3):129-141. Available from: https://journals.ala.org/lrts/article/view/5557/6839

Heidorn P.B. 2011. The emerging role of libraries in data curation and e-science. *Journal of Library Administration* 51(7-8):662-672. DOI: 10.1080/01930826.2011.601269

Hernon P. & Schwartz C. 2015. Descriptions in job ads sometimes equal "Huh?", or "Do you know what you are asking for?" *Library & Information Science Research* 37(3):173-174. DOI: 10.1016/j.lisr.2015.08.001

Holdren, J.P. Increasing Access to the Results of Federally Funded Scientific Research [Internet]. Washington (DC): Office of Science and Technology Policy, Executive Office of the President; Feb 22, 2013 [cited 2016 Nov 13]. Available from https://web.archive.org/web/20161217194429/https://www.whitehouse.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf

Meier J.J. 2010. Are today's science and technology librarians being overtasked? An analysis of job responsibilities in recent advertisements on the ALA JobLIST web site. *Science & Technology Libraries* 29(1-2):165-175. DOI: 10.1080/01942620903579443

National Institutes of Health. 2013. Changes to public access policy compliance efforts apply to all awards with anticipated start dates on or after July 1, 2013 (NOT-OD-13-142) [Internet] Bethesda, MD. Available at https://grants.nih.gov/grants/guide/notice-files/NOT-OD-13-042.html

National Institutes of Health. 2014. NIH Public Access Policy. [Internet] Bethesda, MD. Available at https://publicaccess.nih.gov/

National Institutes of Health. 2016. Clinical and Translational Science Awards. [Internet] Bethesda, MD. Available at https://ctsacentral.org/

National Science Foundation. n.d. Dissemination and Sharing of Research Results. [Internet] Arlington, VA. Available at https://www.nsf.gov/bfa/dias/policy/dmp.jsp

Regazzi J.J. 2013. U.S. academic library spending, staffing and utilization during the Great Recession 2008-2010. *The Journal of Academic Librarianship* 39(3):217-222. DOI: 10.1016/j.acalib.2012.12.002

Si, L., Zhuang, X., Xing, W. & Guo, W. 2013. The cultivation of scientific data specialists: development of LIS education oriented to e-science service requirements. *Library Hi Tech* 31(4):700-724. DOI: 10.1108/LHT-06-2013-0070

Soehner, C., Steeves, C. & Ward, J. E-science and data support services: a study of ARL member institutions [Internet]. Washington (DC): Association of Research Libraries; 2010 [cited 2017 Feb 10]. Available from http://www.arl.org/storage/documents/publications/escience-report-1010.pdf

Tenopir, C., Hughes, D., Allard, S., Frame, M., Birch, B., Baird, L., Sandusky, R., Langseth, M. & Lundeen, A. 2015. Research data services in academic libraries: data intensive roles for the future? *Journal of eScience Librarianship* 4(2):e1085. DOI: 10.7191/jeslib.2015.1085

Tenopir, C., Sandusky, R.J., Allard, S. & Birch, B. 2014. Research data management services in academic research libraries and perceptions of librarians. *Library & Information Science Research* 36(2):84-90. DOI: 10.1016/j.lisr.2013.11.003

Appendix A

Key terms and concepts used to identify job ads with data/informatics responsibilities.

- Computer programming - basic level or understanding (MySQL, SQL, HTML, Python, php)
- Data curation, data preservation, digital preservation
- Data management, data management plans
- Data mining
- Data science, big data, research data
- Digital initiatives
- Electronic record-keeping - electronic laboratory notebooks (ELN), electronic medical records (EMR), electronic health records (EHR)
- Emerging technologies, data visualization

- E-research or e-science
- Geographical information systems (GIS), digital spatial data
- Informatics - including specializations such as bioinformatics, cheminformatics, etc.
- Information systems, database design
- Knowledge management
- Metadata and data standards - terminologies, vocabularies, controlled vocabularies, ontologies, taxonomies
- Open access
- Reference, productivity, or citation management software - RefWorks, EndNote, Zotero, Mendeley, etc.
- Repositories - data, institutional, digital repositories
- Research lifecycle, research data lifecycle
- Research networking software - VIVO, SciVal, Harvard Profiles, etc.
- Research output measurement or assessment - bibliometrics, alternative metrics, ORCiD, etc.
- Scholarly communication
- Semantic web
- UMLS, SNOMED CT

Previous Contents Next