Technical skills for new digital librarians

Elías Tzoc and John Millard

I. Introduction

Over the last decade, libraries, archives and museums have made a major contribution to the creation of historical digital collections which subsequently provide immediate access to primary source materials that might otherwise be unavailable to researchers, scholars, and the general public. A typical workflow for a digital collection project includes processes such as: digitization, conversion and loading data (often in batch), exporting and parsing metadata, and designing web sites.

Two main benefits of digital collections are:

(1) access – whereby institutions can provide multiple and simultaneous users with remote access to a variety of digital objects, including photographs, manuscripts, books, etc.; and
(2) preservation – whereby a digital copy can help preserve the original objects.

Because of the constant changes in technology, librarians working on digital collection projects need to constantly evaluate their access and preservation practices.

This article aims to analyze the current technical skills being sought for digital librarian positions by examining the required and preferred qualifications listed in 43 position announcements posted in 2010, and to explore how well topics – offered by seven major Library and Information Science (LIS) programs in 2010 – have matched these qualifications.

II. Literature review

In an ever-changing technology landscape, the capacity to learn constantly and quickly is more relevant than ever, and is the primary motivation of this study. Similar studies to this one exist in the literature, both as planning tools for hiring administrators on what skills are likely to be needed as libraries change and evolve, or as advice for those new librarians seeking positions in these evolving roles. However, not many have explored the special case of positions focused on digital library development. In 2004, Marion (2000) conducted an exploratory study that analyzed 250 online academic librarian employment ads posted during 2000 to determine current requirements for technologically oriented jobs. Marion organized the results into different categories, and two of those categories seem relevant to this study: programming languages and web site creation.

Perhaps the best attempt at addressing digital library competencies, Choi and Rasmussen (2006) conducted a survey for their 2006 article: "What is needed to educate future digital librarians" to identify their skills and to detect possible gaps in their training. The authors concluded "LIS education needs to pay attention to [. . .] integration of practical skills and experience with digital collection management and digital technologies into curricula."

In 2007, Tammaro (2007) analyzed the trends for digital library education in Europe. Although the focus of this work was about a "curriculum for digital librarians," the first set of competencies (information architecture, information retrieval, web-publishing, database theory, networking, human computer interaction, evaluation of information systems, and technical troubleshooting skills) are directly related to what this study plans to examine in the list of technical courses currently offered in library programs. Most recently, in the Spring of 2009, Mathews and Pardue (2009) conducted a five-month study to analyze the IT skills that employers were looking for when posting job ads in ALA's online JobLIST.

III. Methodology

Based in part on this literature review, a twofold data collection methodology was developed that compared the required and desired technical skills as expressed in position announcements and the skills currently being taught in major LIS programs. A common set of categories was developed to account for variations in wording and specific implementations of a technology.

3.1 Data collection

Analysis of required and preferred qualifications: to ensure the most current data, a 12-month period was selected for analysis, from January to December 2010. Data were collected from five sources: ALA JobLIST, job opportunities from EDUCAUSE, LISJobs.com, and three library schools' career sites. Position announcements were limited to those for digital collection/digital library-related positions.

Analysis of technical courses: included an analysis of technical courses offered in 2010 at selected library programs listed among the top schools as identified by the 2009 *US News and World Report* ranking. An initial list included the top five schools specializing in archives and preservation, information systems, and digital librarianship. Duplicates were removed from the combined list resulting in a final list of seven target schools (*US News and World Report*, 2009).

3.2 Data pre-evaluation

One challenge in gathering historical data for position announcements is that most job banks only keep the data for a given period of time – usually from 60 to 90 days. The EDUCAUSE's job opportunities web site was very helpful as it provides access to past data. The availability of dedicated career services at some library schools was helpful. Position announcements posted in ALA JobLIST expire after 90 days. However, ALA was able to provide archived data for this study.

Another challenge in the pre-evaluation process was the discovery of new position titles with similar requirements to those found for positions specifically for digital collection projects. Extra care was given to ensure that the job description and responsibilities for these new positions were within the scope of the study. The resulting pool of position announcements to be analyzed included 43 distinct positions working exclusively on digital library or digital collection activities. Table I summarizes the source of the position announcements analyzed.

An interesting finding from the pre-evaluation process was that 46 percent of positions preferred a bachelor's degree in computer science, information management or a related field. This may be a confirmation of the ongoing need to recruit new librarians with those types of degrees. The data were normalized to a common format to account for differences in descriptive style and length of entry. The statistical reports were generated using features in MS Excel such as pivottable report.

Table II represents 21 unique position titles. The differences in position titles may represent the diversity of understanding of where these kinds of positions fall within the organizational framework or workflow of the hiring institution.

Data on technical courses offered by the selected LIS programs were collected from descriptions of course offerings during the spring and fall semesters of 2010. Data collected included course title and description. Courses were assigned a general category based on topic to allow for more meaningful comparison. Courses without enough contextual information in the description field were not included in the sample of 46 courses. Courses with a direct link to syllabus were extremely helpful, especially where faculty listed specific class projects. Table III provides a summary of the 17 technical courses offered in 2010 in the sample of seven schools with a specialization in archives and preservation, information systems, or digital librarianship.

Table I.
Summary of job ads grouped by source

Source	Total	Percentages
ALA JobLIST	17	40
EDUCAUSE Jobs	9	21
iSchool-Career-Services	10	23
LISJobs	7	16
Total	43	100

Table II.
Position titles represented in the sample

Position title	Total	Percentage
Data librarian	1	2
Digital archivist	2	5
Digital collection/metadata librarian	1	2
Digital collections librarian	3	7
Digital collections metadata librarian	1	2
Digital collections specialist	1	2
Digital and scholarly communications librarian	3	7
Digital initiatives librarian	2	5
Digital integration librarian	1	2
Digital librarian	3	7
Digital scholarship and services librarian	2	5
Digital services librarian	4	9
Information technology librarian	2	5
Librarian digital services support	1	2
Library digital services manager	3	7
Metadata and digital initiatives developer	1	2
Metadata and digital resources developer	1	2
Systems and digital collections librarian	2	5
Systems development librarian	2	5
Web initiatives librarian	2	5
Web services librarian	5	12
Total	43	100

3.3 Categorization

Position announcements were analyzed for required and desired skills. The resulting list of skills and qualifications were grouped into common broad categories and a frequency distribution was created. The same analysis was then performed on course descriptions using the same common categories. In all, 12 broad categories were identified:

(1) Database design and management – knowledge of and experience with relational database design, deployment and management including proficiency with SQL on both commercial and open source database servers.
(2) Digital collection management – understanding and experience managing all facets of digital collections including overall technical proficiency in the tools and technologies related to digital repositories and digital collection development.
(3) Digital content management systems – experience using or managing common content

Table III.
List of technical courses represented in the study

Course	Total	Percentage
Advanced database management	2	4
Advanced web technology/presentation	3	7
Content management system	1	2
Data interoperability	2	4
Development of web applications	5	11
Digital library technology/software	3	7
Encoded archival description	1	2
Information architecture for the web	3	7
Information visualization	2	4
Introduction to database management	6	13
Introduction to web design	4	9
Introduction to web programming	6	13
Mobile application development	1	2
Usability	1	2
Web 2.0	2	4
Web archiving	1	2
Digital preservation and access	3	7
Total	46	100

management systems including digital asset management systems. Specific examples include DSpace, CONTENTdm, Drupal, Luna Insight, DLXS, WordPress, and Omeka.

(4) Digital conversion – knowledge of the conversion of analog materials to digital formats in a library context. It includes knowledge of digitizing equipment and best practices, archival file formats, and reformatting audio and video materials.

(5) Digital preservation – knowledge and experience with preservation of both analog and digital materials and ability to manage the ongoing preservation of digital collection content.

(6) Metadata and cataloging standards – familiarity with emerging or established metadata and cataloging standards including Dublin Core, EAD, TEI, FRBR and RDA as well as controlled vocabularies.

(7) Programming: JAVA, C, C++ – familiarity with more formal development languages like JAVA, C, C++, etc.

(8) Programming: scripting languages – familiarity or experience with less formal procedural scripting languages like PHP, Perl, JavaScript, Ruby and Python.

(9) Systems and network administration and desktop support – experience administering Unix/Linux/Windows servers, designing and managing networks, and providing desktop configuration and support for both Mac and Windows PCs.

(10) Web application development – experience using common scripting languages, relational databases in the creation of dynamic date-driven web sites and applications.

(11) Web design and web standards – understanding of web design standards including HTML, XHTML, CSS, interface design and common software platforms in web development like Dreamweaver.

(12) XML and related standards – primarily, experience with XML and related technologies like XSLT and XML databases.

IV Results

Each position description was tagged with all applicable categories. Once tagged, a simple distribution was developed showing the percentage of position descriptions requiring or desiring qualities in each category. As shown in Figure 1, the top six categories were web design and web standards (72 percent of the sample of 43 job ads included this category), digital collection management and programming with scripting languages (64 percent), digital content management systems, metadata/cataloging standards, and XML/XSLT (48 percent). The least expressed category in this study was digital preservation (8 percent).

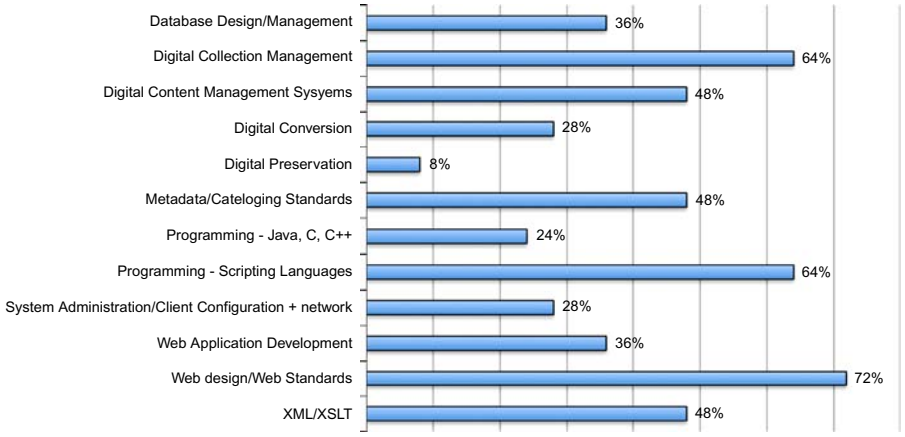
The same methodology applied to technical courses in LIS programs. Figure 2 shows the following: the top four categories represented were web design and web standards (30 percent of the sample of 46 technical courses offered this topic), database design and management (27 percent), web application development (24 percent), and digital collection management (20 percent). The three top popular topics in web design and web standards courses were: web navigation, information visualization, and user-centered design. For database design and management, four schools offered two courses: introductory and advanced database mainly focused on SQL and web database applications.

For comparison purposes, the two sets of data (job ads and courses) were re-calculated in a scale of 1-100. Figure 3 shows the correlation of technical skills' categories between position descriptions and LIS courses.

Generally, skills taught in LIS courses closely matched those being sought in position announcements with some notable exceptions. As shown in Figure 3, digital preservation was not a highly sought after qualification with only 2 percent of the 43 positions requiring it. Conversely, 7 percent of courses included this topic. LIS curriculum may be ahead of common practice in the field here or employers may be assuming that digital preservation is a standard skill set common in all LIS program graduates. Further study is needed to determine the cause of this disparity.

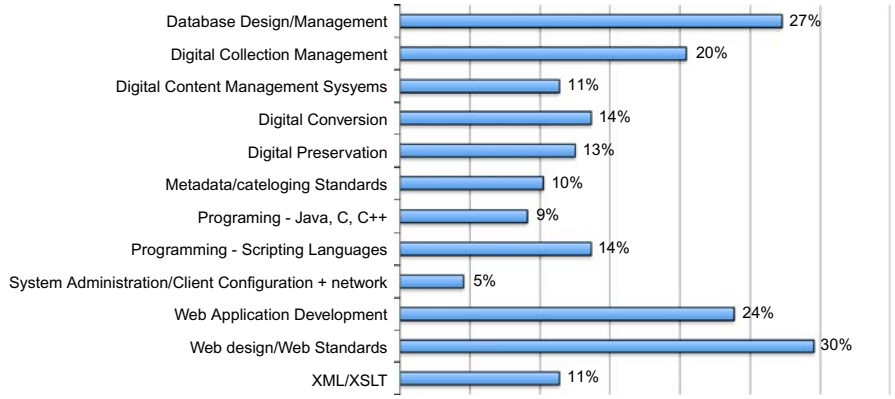
More significant disparities existed in the distribution of programming – scripting languages, digital content management systems, metadata and cataloguing standards, and XML/XSLT skills. It is clear that these skills are perceived as core competencies for these types of positions. LIS curriculum

Figure 1. *Percentages of technical skills sought by employers*



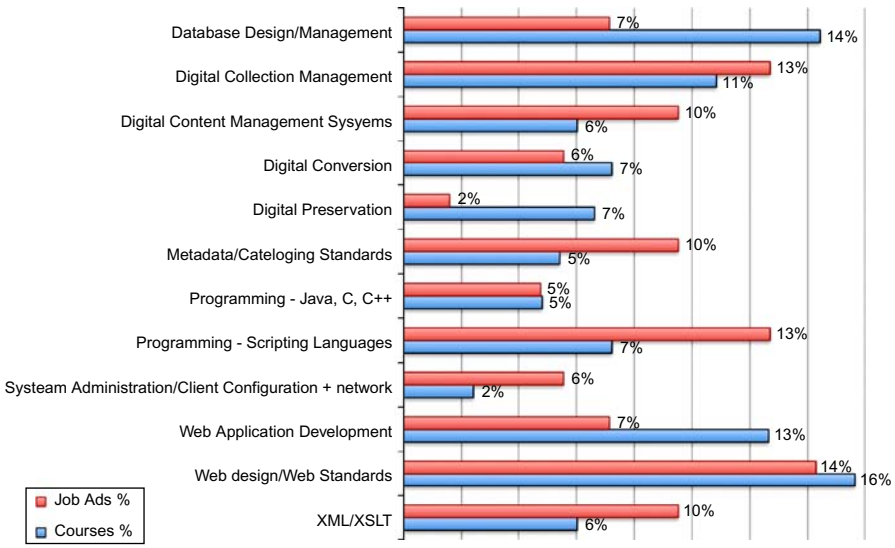
Note: % = n/43

Figure 2. *Percentages of technical courses in LIS programs*



Note: % = n/46

Figure 3. *Categories comparison: position descriptions vs LIS courses*



is clearly leading in three areas: database design and management, web application development, and web design and web standards.

A certain amount of disparity in results can be expected due to the differing missions between hiring institutions and LIS programs. For example, it is possible that LIS curricula look further ahead to anticipate future developments in the field. Indeed, our review of the literature review includes attempts to determine the best curriculum to train future students. In addition, the analysis of topics taught revealed several topics taught that were not represented in the position announcements at all. Notable among these were courses covering open source software evaluation, assessment, and information visualization.

V Conclusions

One conclusion of this analysis is that current students as well as practicing librarians need to seek out additional non-curricular opportunities to build competency in the technical areas represented in this study if they are or expect to be marketable. Fortunately, the areas where the greatest disparity exists are also areas where significant opportunities for independent learning are available. For instance, one could begin to gain experience individually by setting up a local and general-purpose web server (sandbox) with a basic LAMP/WAMP configuration. From here, one can install common open source content management systems such as Drupal, Word press, DSpace, or Omeka and begin creating small collections for testing and experimentation. For more specific or advanced tutorials, free online training sites such as W3Schools.com or subscription based such as Lynda.com can be valuable resources as well. Additionally, new digital librarians may also join technical groups such as Code4Lib, where members are constantly exchanging ideas on top technical trends and development in the library community. A good example is a May 2011 discussion in the Code4Lib listserv about what library technologists "would like to learn", which provides a useful list of technologies currently in demand in

the library world (http://tinyurl.com/code4lib-time-to-learn).

Another conclusion is that students in library programs should be encouraged to apply or volunteer for internship opportunities or work on technical capstone projects, as these opportunities will help them acquire real-life experience. In fact, two to three years of experience was indicated as a preferred qualification in more than 40 percent of job ads analyzed in this study. Having the opportunity to work on a real project will help new librarians acquire the skills for which employers are looking. And for those new librarians already working on digital collection/library programs, a personal interest in continuously learning and improving their technical skills will be essential for their professional development.

Finally, during the data collection of this study, the authors found multiple positions for other technical/digital areas of academic librarianship – but with similar technical requirements. One of the new titles found repeatedly was Emerging Technologies Librarian where the main responsibilities include the exploration and creation of mobile applications as well as implementation of APIs and mashup tools such as Google Maps for geotagging. A further study of the overlap of responsibilities and expectations for digital librarians in other areas of academic libraries will be useful.

ACKNOWLEDGEMENTS

The authors would like to recognize the contribution of David M. Connolly, Classified Ads Coordinator from the American Library Association, in providing archived data for this publication.

REFERENCES

Choi, Y. and Rasmussen, E. (2006), "What is needed to educate future digital librarians", *D-Lib Magazine*, Vol. 12, available at: www.dlib.org/dlib/septembelr06/choi/09choi.html (accessed 12 January 2011).

Marion, L. (2000), "Digital librarian, cybrarian, or librarian with specialized skills: who will staff digital libraries?", Association of College & Research Libraries Publication, available at: www.ala.org/ala/mgrps/divs/acrl/events/pdf/marion.pdf (accessed 23 March 2011).

Mathews, J. and Pardue, H. (2009), "The presence of it skill sets in librarian position announcements", *College & Research Libraries*, Vol. 70 No. 3, pp. 250-7.

Tammaro, A. (2007), "A curriculum for digital librarians: a reflection on the European debate", *New Library World*, Vol. 108 Nos 5/6, pp. 229-46.

US News and World Report (2009), "Library and information studies ranked report", *US News and World Report*, available at: http://grad-schools.usnews.rankingsandreviews.com/best-graduate-schools/top-library-information-science-programs (accessed 17 August 2010).

Elías Tzoc (*tzoce@muohio.edu*) is the Digital Initiatives Librarian at Miami University Libraries, Oxford, Ohio, USA.

John Millard is Head of Digital Initiatives at Miami University Libraries, Oxford, Ohio, USA.