Education for eScience Professionals:
Integrating Data Curation and Cyberinfrastructure

Youngseek Kim, Benjamin K. Addom, Jeffrey M. Stanton,

School of Information Studies, Syracuse University

Abstract

Large, collaboratively managed datasets have become essential to many scientific and engineering endeavors, and their management has increased the need for "eScience professionals" who solve large scale information management problems for researchers and engineers. This paper considers the dimensions of work, worker, and workplace, including the knowledge, skills, and abilities needed for eScience professionals. We used focus groups and interviews to explore the needs of scientific researchers and how these needs may translate into curricular and program development choices. A cohort of five masters students also worked in targeted internship settings and completed internship logs. We organized this evidence into a job analysis that can be used for curriculum and program development at schools of library and information science.[1]

Introduction

The problems arising from collecting, organizing, indexing, archiving, and sharing large datasets have increased the need for information professionals who offer a mixture of science or engineering knowledge together with the capabilities taught in a range of educational programs in Library and Information Science (LIS). These emerging "eScience professionals" may serve in a new professional area of librarianship that solves large scale information management problems for researchers and engineers with innovative tools and techniques. In an informal job market analysis for professional positions in this area, we found 208 eScience professional positions from two job search websites including HigherEdJobs.com and Monster.com during a one month period (from February 26th to March 27th, 2010). The eScience professional positions appear with diverse job titles including data analyst, research analyst, information specialist and data specialist.

In this paper we report on a program of research in which we interviewed researchers and sent a small cadre of information professionals-in-training on guided internships in scientific laboratories. We have analyzed the resulting data to triangulate on the areas of knowledge and skill that eScience professionals must possess, and from these knowledge and skill areas have offered suggestions and possibilities for curriculum and program development.

Background

In 2001, John Taylor, the Director General of Research Councils at the Office of Science and Technology in Great Britain, articulated a vision for large scale scientific collaboration: "e-Science is about global collaboration in key areas of science and the next generation of infrastructure that will enable it" (Hey & Trefethen, 2003a). Shortly thereafter, in the U.S., a National Science Foundation panel, headed by University of Michigan School of Information dean Dan Atkins, described similar sentiments, but expanded the scope beyond science and into engineering and industrial research and development through a newer term: "Cyberinfrastructure", which refers to infrastructure based upon distributed computer, information and communication technologies (Atkins et al., 2003). In both cases, the vision included recognition that a "data deluge" would arise as an inevitable side effect of eScience and cyberinfrastructure (Hey & Trefethen, 2003b). The availability of large datasets and the capability of productively sharing these datasets among international teams of researchers was framed as, in effect, both a central goal and a critical challenge of eScience and cyberinfrastructure.

As eScience and cyberinfrastructure, including data curation and information management, become critical in the advancement of science and engineering, educational programs designed to train eScience information professionals could provide a professional path into a powerful and valuable set of roles within a variety of research enterprises. Several authors have begun to explore possible dimensions of such programs. For example, a number of researchers have worked on digital curation curricula (Gordon, 2009; Grace, Anderson & Lee, 2009; Waters & Allen, 2009). Pomerantz et al. (2009) described a striking degree of overlap between a library science curriculum designed for teaching about digital libraries versus a curriculum focusing on digital curation (specifically, the DigCCurr project). Renear et al. (2009)

described how a library science curriculum suitable for digital curation could be beneficially extended to serve researchers in the humanities. Yakel, Conway & Krause (2009) highlighted their efforts to provide students with targeted internships, including work at research organizations such as the Smithsonian Institution and the Inter-University Consortium for Political and Social Research.

In contrast to approaches described above, which build curriculum for information professionals based on digital librarianship and curation, some have advocated an educational approach that focuses on cyberinfrastructure and its consequent challenges in grid computing, service oriented architectures, simulation, virtualization, sensor networks, and collaboration tools. For example, given the emerging importance of interdisciplinary research areas such as bioinformatics, there is an evident need to educate individuals who have domain knowledge in a discipline such as biology as well as mastery of techniques such as gene sequence analysis, protein folding simulations, and computational modelling of regulatory pathways (W. Wang, 2009). Relatedly, considerable effort has been spent on curriculum development issues for the education of biological information specialists (Heidorn, Palmer, Cragin & Smith, 2007; Heidorn, Palmer & Wright, 2007; Palmer, Heidorn, Wright & Cragin, 2007). It is very interesting to note, however, that there appears to be minimal détente between curation educators and infrastructure educators. In a 293-page compendium of papers entitled: "Transform Science: Computational Education for Scientists" (Xu, 2009) the word "archive" is mentioned just once, "metadata" just once, and "curation" not at all.

Rather than choose between these positions, we began from the premise that some mix of skills and knowledge from areas closely identified with librarianship together with capabilities for making use of cyberinfrastructure tools might best serve students who choose to enter this specialized area. This notion of a mixture of information skills and science skills is consistent with the approach previously taken by Heidorn, Palmer and Wright (2007) in their analysis of the role of biological information specialists. To explore the nature and balance of this mix, we undertook a set of data collections that amounted to a job analysis of the emerging position of eScience professional. Job analysis – the systematic investigation of work roles and worker qualifications – has been used repeatedly by librarianship researchers to help understand the changing needs and demands of work roles in libraries (e.g., Ricking & Booth, 1974). For decades, job analysis researchers and practitioners have used two complementary strategies for understanding the nature of specific work roles: analyzing the work itself and analyzing the qualifications of the worker (Stetz, Button & Porr, 2009). Worker-oriented job analysis methods such as the Job Element Method, the Position Analysis Questionnaire, and Cognitive Task Analysis, focus on the worker's qualifications in conducting tasks in a job position. Work-oriented job analysis methods such as Functional Job Analysis, Task Inventories, and the Critical Incident Technique focus on inventorying the major work duties for a given position. In the present study, we used Fine and Cronshaw's (1999) job analysis framework, which hybridizes the work and worker approaches and which also takes the organizational context into account.

Method

Overview

Five interviews and five focus groups were conducted to collect the job requirements of eScience professionals with respect to the work itself, worker qualifications, and work organizations. Participants for the interviews and focus groups included eight laboratory directors, seven researchers in science and engineering research centers. We asked these participants to describe the duties of eScience professionals; the characteristics of their work environments, such as organizational structure and hierarchy; the required characteristics of workers, including knowledge, skill, ability and attitudes; and any specific tools, equipment or materials that workers needed to use or master. Note that the focus groups were conducted with curriculum development in mind, modelling a common strategy for eliciting task requirements (Witkin & Altschuld, 1995).

Subsequently, we placed five eScience professional masters students in summer internships. During the internships we elicited a variety of task, skill and knowledge data from the students, first at one week intervals for about a month and subsequently at two week intervals for roughly two more months. Our students worked closely with researchers and kept detailed logs of their activities. At the end of their internships, they completed an exit questionnaire that rated a range of tasks in both technical and content areas in terms of frequency, importance and change in capabilities.

Data Analysis

The interviews and focus groups were tape-recorded and transcribed. Each interview took 25-35 minutes and each focus group took one to two hours. After we transcribed all of the interviews and focus groups, we imported them into "QDA Miner," a software application optimized for textual data analysis. To conduct the data analysis, we developed our own coding scheme by using both deductive and inductive approaches. We used Fine and Cronshaw's (1999) job analysis framework and its components in order to create a general data analysis scheme, and we also used an inductive approach to create more specific codes within each major category of Fine and Cronshaw's original framework.

An initial coder processed the transcripts and applied codes to 755 relevant utterances in the interviews and focus groups. A second coder reviewed the initial coding to identify areas of disagreement in the coding. The two coders disagreed on approximately 7% of the codes (53 out of 755). Following discussion and clarification of coding rules, the coders reduced their disagreements to less than 1% of applied codes (8 out of 755). The first coder then re-applied the updated scheme to the whole dataset after which the second coder independently coded a random sample of approximately 75 utterances. The final level of inter-coder agreement was 87%. Subsequently, we analyzed summer internship logs and exit questionnaires (1473 total utterances) using the same coding scheme originally developed for interviews and focus groups. We found that the original coding scheme was applicable in summer internship logs, with two codes added to address novel content. Note that in the presentation below we have intentionally avoided presenting a quantitative analysis of code occurrences, because the focus groups were conducted in a way that sought to obtain consensus, and therefore contained considerable repetition.

Results

Job Analysis Results

Based on our content coding of 755 utterances from interviews and focus groups, we observed six major duties with respect to "data," including collecting primary data (cleaning and checking data, collecting original data, and understanding data needs), collecting secondary data (such as previous literature or public/commercial data sets), storing data (creating databases, managing metadata, and storing data), managing data (cleaning, annotating, managing, maintaining, and future planning), analyzing data (statistical analysis, processing scripts), and presenting data (helping researchers to access data, posting data for wide access, dealing with data ownership, and writing about data).

Along with the data related duties we also identified a number of major duties for the eScience professionals in terms of working with "people": locating collaboration opportunities, communicating with others, enabling collaborations and organizing teams, analyzing researchers' technology needs, coordinating between researchers and information technology experts (e.g., with technology requirements and specifications), ensuring compliance, and training researchers and others in using technologies. Finally, we found several major duties that mainly pertained to the use of "things" – primarily computers and software: investigating technology solutions, recommending technology solutions (by comparing technologies), implementing IT for researchers (installing operating systems, installing software applications, managing collaborative technologies, and configuring systems by using scripting), maintaining and managing the technologies (administering systems, maintaining tools/technologies, and facilitating IT usage), preparing, compiling, and managing documents, and managing budgets and project processes.

Fine and Cronshaw's job analysis framework ([1999](#)) also calls for analysis of worker characteristics required for effective performance on the job, including knowledge areas, skills and abilities. Knowledge refers to a body of information that must be memorized or mastered. Participants identified several major knowledge areas including knowledge of databases, of terminology and methods in scientific subject domain area (e.g., physics), of information technology, and of programming or scripting languages. Next, skills are acquired competencies subject to education, training and improvement with practice. We identified eight different skills: administrative skills, communication skills, database management skills, programming and scripting skills, project management skills, research skills, system administration skills and general computer skills. Finally, ability refers to one or more intrinsic talents an individual possesses. Abilities represent an individual's potentialities in a given area. Abilities may overlap with skills in that ability signifies the capacity to perform in a certain class of tasks (e.g., perceptual acuity), whereas a skill represents the extent to which a task has been mastered through practice. Participants described three areas of ability: ability to work well in a team environment, ability to quickly learn new material and ability to communicate with others.

Participants in our focus groups and interviews also commented on the education, experience and tools needed for effective performance in the eScience professional position. Participants recommended interdisciplinary education in science or engineering together with training in information technology. Participants also recommended coursework in database development, data management, project management, research methods, statistics and programming or scripting classes. The tools that eScience professionals need include collaboration software, data sharing applications, database systems, project management software, qualitative data analysis software, security technology, web applications, operating systems, and servers.

Finally, we also asked our focus group and interview participants about the organizational environment in which eScience professionals conduct their work. Participants suggested that these professionals usually work in science and engineering research centers in academia or industry. In some cases, they may also work in social science and policy research centers. Organizational missions in these environments focus on developing scientific findings by analyzing large amounts of data and making those data accessible to other researchers. Participants saw collaboration with other relevant research centers as a key goal; organizations fulfill this goal by making the data as accessible as possible, releasing government information to the public, publishing research findings and maintaining shared data repositories. Participants also commented on the major problems their research centers have encountered: overloading by huge data sets, managing databases, dealing with new technologies, collaborating with distributed teams and the complexity of their data problems.

Summer Internship Analysis Results

Based on our content coding of 1473 descriptive utterances from the summer internship logs, we obtained general information regarding the work duties that the five students performed and the various organizational environments in which they worked. Students primarily worked with researchers and information technology professionals throughout their internships. Three of the five students also worked with professional colleagues whose work overlapped strongly with eScience, including science librarians, a metadata librarian, and a project information manager. In addition, we found that the students' internship organizations faced similar problems as those identified by interview and focus group participants: huge data sets, the challenges of managing multiple large databases, constant learning needed to deal with new technologies, and the importance of collaborating with geographically distributed teams.

With respect to Fine and Cronshaw's work duty categories ([1999](#)) of data, people, and things, the results from the task analysis of internship logs were strikingly similar to those obtained from the interviews and focus groups. Three of five students mainly worked on data related tasks: data collection (especially secondary data), storage, management and presentation. Students defined and maintained metadata, and spent considerable time on data integrity, cleaning data and annotating datasets. Students created graphics, designed interfaces, prepared presentations. In contrast to the focus groups and interviews, these students had few chances either to collect primary data or to conduct statistical analysis.

All five students had extensive responsibilities for communicating with other people during their internships, particularly with respect to obtaining, understanding, and translating the data needs of researchers. Relatedly, students reported that they conducted user training activities. Two students had responsibilities for facilitating geographically distributed communications. Likewise, two students were involved with administrative work such as managing project processes and budgets, preparing meeting or conference materials and planning travel.

Finally, all five students had responsibilities for investigating technology solutions for various data management, storage and analysis problems. Students recommended possible technology solutions and three of the five students were further involved in both technology acquisition and implementation processes. Further, these same three students participated in the actual technology implementation processes by programming, scripting and configuring various technology components. Lastly, students also managed and maintained some technologies including websites, various hardware and software, and sometimes servers.

In addition to examining the internship logs for clues about the work duties and organizational settings of the internship students, we asked the students to report on the frequency and criticality of the tasks that they had performed. We elicited these ratings following completion of each student's internship. Figure 1 displays the results.
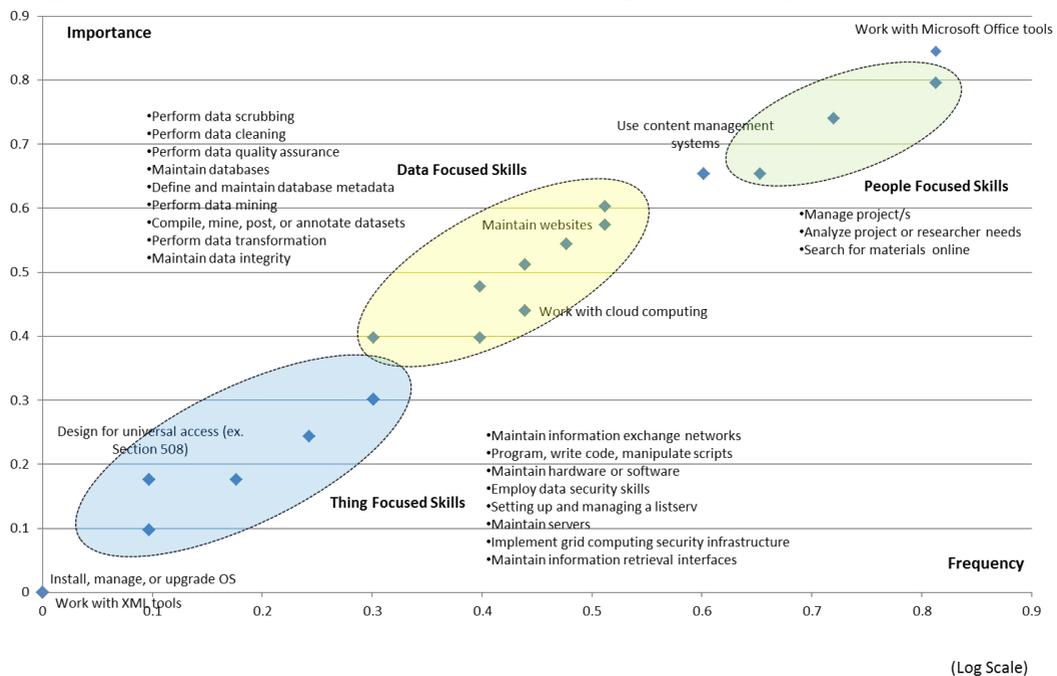

Figure 1. Student Ratings of Relative Importance and Frequency of Internship Tasks.

The shaded areas in Figure 1 represent groupings of the students' internship tasks into the three major categories of data, people and things. Notably, a large set of tasks in the data area – from maintaining data integrity to defining metadata – received substantial importance in the students' ratings. In the "things" area, working with content management tools and office productivity software obtained the greatest importance. With respect to people, managing projects and analyzing project and researcher needs obtained the greatest importance.

Internship Exit Interviews

All five students reported that domain area knowledge in a scientific discipline was critical during their internship. Students also found that they needed technical knowledge including database, programming (or scripting), and general information technology knowledge. They also needed knowledge in research methods and research ethics. Regarding skill sets needed for the internships, students mainly reported a need for skills in communication, database management, programming/scripting, project management, research, system administration, and general computer skills. Finally, students reported numerous software tools used during their internships in the areas of database management, citation organization, data sharing, project management and collaboration.

Discussion

Overview of Findings

We have presented a variety of qualitative analyses of focus groups and interviews with lab directors and researchers, together with some descriptive analysis of rating data and coded work activity logs provided by five summer internship students. These analyses converged on the importance of a tripartite role for eScience professionals including data curation, communication and cyberinfrastructure. First and foremost, these professionals need to have a range of data curation capabilities – including knowledge, skills, classroom experience and familiarity with infrastructure tools – for managing large, complex, interconnected databases of research data. The specific tasks in this cluster, pertaining to data quality assurance, data integrity maintenance, data access and metadata, are closely aligned with emerging ideas of digital data curation roles (Baker & Yarmey, 2009; Becla & Lim, 2008; Hazeri, Martin & Sarrafzadeh, 2009).

Second, the eScience Professional, as embodied in the comments of lab directors and the experiences of our internship students, seems to play critical roles in bridging communications between the research community and the IT infrastructure community. Throughout our data, communication skills repeatedly arose as an essential area of capability for these professionals; communication in service of fathoming researcher needs, managing projects, and facilitating collaborations among the various communities involved in the research process. These various, diverse collaborations seemed particularly important given the complexity of the data curation involved in the scientists' research processes. Unlike the research of decades ago, contemporary research requires the concerted efforts of many distributed teams that are involved in data collection, data preparation, data analysis and data archiving. Without a sophisticated set of skills for fostering collaboration and effective project management across a network of cooperating sites, our students would have been much less valuable to their internship hosts (Kinkus, 2007; Promis, 2008).

Finally, each of our internship students functioned in one or another paraprofessional role, swinging the pendulum either toward the scientist – in which they were expected to conduct tasks such as literature review, secondary data collection, data cleaning and data mining – or towards the IT professional. In this latter role, students had responsibilities such as working with web content management systems, cloud computing, grid computing and even some light duties in scripting or

programming. As such, varying amounts of expertise in both a science discipline and technology or cyberinfrastructure were needed, depending upon the specific job role the student fulfilled (Aloisio & Fiore, 2009; Hey & Trefethen, 2003b; von Laszewski, 2009; S. Wang, Liu, Wilkins-Diehr & Martin, 2009).

Curricular Considerations

Our students entered their internships with a range of different courses under their belts, including core library science courses for some and more technologically focused coursework for others. In all cases, though, students had taken a course on databases and a course on scientific data management, both of which proved essential in preparing them for their internship experiences. Additionally, although only one of our students had taken a project management course prior to their internship, the extent to which project management skills seemed to figure prominently in all of the students' internships suggested that project management ought to be a required course for anyone seeking to become an eScience professional. In the same vein, we heard from both scientists and students that some capability for scripting or programming was highly worthwhile.

We present below a top ten list of recommended courses that may have greatest value for individuals who aspire into eScience professional roles:

- Digital data curation, optimally in a course specialized toward the curation of large scientific or engineering datasets;
- Database design and management, focusing on large scale relational databases;
- Project management, including project planning and budgeting;
- Essentials of scientific research, including literature review, study design, and descriptive statistical analysis;
- Overview of cyberinfrastructure, including cloud and grid computing;
- Geographically distributed collaboration, with a judicious division of time between the human issues and the technological issues;
- Web content management and web interaction design;
- Scripting or practical introductory programming;
- Data mining, with a focus on either quantitative data for the natural sciences or mixed data types for the social sciences and humanities;
- Information system management and server administration, including general IT and computer knowledge.

Many of these topics are represented within collections of courses commonly available in library science, information science and information studies programs. Therefore, there are a number of ways of dividing these topics into courses, so this list should be considered as input into curricular deliberations rather than a fixed course list. Additionally, note that students' capability to excel in the topics described above probably requires a base level of exposure to essential concepts and skills in information science that a student would typically obtain from core coursework.

Limitations

Although our study drew upon a rich corpus of qualitative data to understand the work activities and environments of eScience professionals, it is possible that some of our findings are idiosyncratic to the small group of students and professionals involved in our study. Notably, all of our informants – lab directors, researchers, and students alike – were involved in organizations in the non-commercial sector, primarily in academia. We believe, however, that there is great potential for employment of eScience professionals on the front lines of commercial activities, particularly those in the pharmaceutical, medical and biosciences sectors. Many firms in these sectors that have active research and development activities will find a need to employ individuals with the range of skills and knowledge in the areas of data curation and cyberinfrastructure.

In a similar vein, because the roles of eScience professionals are still emerging and evolving in the workplace, it is likely that the definition of a job in the eScience area is still a moving target. For these reasons, the work we have reported here should be construed as an initial leg in an ongoing process of triangulation. As the number of information professionals employed in eScience grows, it will be important to periodically re-examine the range of duties, skills and educational preparation involved in the work.

Conclusion

The unique value of eScience professionals is that their work enhances the progress of research in science and engineering endeavors by managing the deluge of scientific information. Sophisticated, professional information management allows scientists to do the best possible science and IT professionals to create the most reliable, cost effective and capable infrastructure. Problems in both of these domains are sufficiently complex that having the "bridge function" – the capability of translating the information needs of scientists into cyberinfrastructure tools – is indispensable. The present research has contributed to an improved understanding of the job requirements for professionals who occupy this bridge role, but there is much work left to be accomplished. While many library scholars have studied employment trends in various segments of the library job market (e.g., Beagle, 1999; Beile & Adams, 2000; Borko, 1984; Boyd, 2008; Kieserman, 2008; Lai, 2005), we believe that the eScience area regarding data curation and cyberinfrastructure deserves more attention. As a form of "embedded librarianship" where information professionals serve right in the midst of the research and development activity alongside scientists and technology specialists, eScience job roles may embody the new wave of librarianship where the four walls of the library are less important than the unique capabilities that librarians bring to hard problems.

In addition, it is important to remember applications of this unique skill set are not limited to the natural sciences. Burton and Appleford (2009), Renear et al. (2009), and others have aptly documented the possibilities for applying cyberinfrastructure to research in the arts, humanities, and social sciences. As well, large scale data management problems are not limited to research and development endeavors – government, corporate, and educational sectors have also begun to experience their own flavors of the eScience data problem. On this basis, we expect that significant demand will arise for individuals with eScience professional skills in terms of data

curation and cyberinfrastructure, that numerous other institutions of higher education will need to join the process of educating them, and that a significantly expanded supply of students to join these programs will be required. In short, we conclude that serving the education and training needs of eScience information professionals can be a promising curricular and program focus in both the LIS and data curation educations over the coming decade.

Acknowledgements

References

Aloisio, G., & Fiore, S. (2009). Towards exascale distributed data management. *International Journal of High Performance Computing Applications*, *23, (4)*. Retrieved February 1, 2011, from http://hpc.sagepub.com/content/23/4/398.

National Science Foundation (2003). Revolutionizing science and engineering through cyberinfrastructure. *Report of the National Science Foundation Blue-Ribbon Advisory Panel on cyberinfrastructure*. Retrieved February 1, 2011, from http://www.nsf.gov/od/oci/reports/atkins.pdf.

Baker, K., & Yarmey, L. (2009). Data stewardship: Environmental data curation and a web-of-repositories. *International Journal of Digital Curation*, *4, (2)*.

Beagle, D. (1999). Conceptualizing an Information Commons. *Journal of Academic Librarianship*, *25, (2)*. Retrieved February 1, 2011, from http://linkinghub.elsevier.com/retrieve/pii/S0099133399800032.

Becla, J., & Lim, K. (2008). Report from the first workshop on extremely large databases. *Data Science Journal*, *7*. Retrieved February 1, 2011, from http://joi.jlc.jst.go.jp/JST.JSTAGE/dsj/7.1.

Beile, P. M., & Adams, M. M. (2000). Other duties as assigned: Emerging trends in the academic library job market. *College & Research Libraries*, *61, (4)*. Retrieved February 1, 2011, from http://crl.acrl.org/content/61/4/336.

Borko, H. (1984). Trends in library and information science education. *Journal of the American Society for Information Science*, *35, (3)* Retrieved February 1, 2011, from http://onlinelibrary.wiley.com/doi/10.1002/asi.4630350312/abstract.

Boyd, R. (2008). Staffing the Commons: Job analysis in the context of an Information Commons. *Library Hi Tech*, *26, (2)*. Retrieved February 1, 2011, from http://www.emeraldinsight.com/journals.htm?articleid=1729332.

Burton, O. V., & Appleford, S. (2009). Cyberinfrastructure for the Humanities, Arts, and Social Sciences. *Research Bulletin of the Educause Center for Applied Research, (1)*. Retrieved February 1, 2011, from http://www.educause.edu/ECAR/CyberinfrastructurefortheHuman/163575.

Fine, S., & Cronshaw, S. (1999). *Functional job analysis: A foundation for human resources management.* Mahwah, New Jersey: Lawrence Erlbaum.

Gordon, K.M. (2009). A residential data curation internship: Opportunities and challenges. *Proceedings of DigCCurr2009 International Symposium on Digital Curation.* Chapel Hill, NC, USA.

Grace, S., Anderson, S., & Lee, C.A. (2009). Skills for significant properties: Debating pragmatics and philosophy in an area of digital curation. *Proceedings of DigCCurr2009 International Symposium on Digital Curation.* Chapel Hill, NC, USA.

Hazeri, A., Martin, B., & Sarrafzadeh, M. (2009). Integration of knowledge management with the library and information science curriculum: Some professional perspectives. *Journal of Education for Library and Information Science, 50, (3)*. Retrieved February 1, 2011, from http://jelis.org/2009-volume-50/volume-50-number-3/.

Heidorn, P.B., Palmer, C.L., Cragin, M.H., & Smith, L.C. (2007). Data curation education and biological information specialists. *Proceedings of DigCCurr2007 International Symposium on Digital Curation.* Chapel Hill, NC, USA.

Heidorn, P.B., Palmer, C.L., & Wright, D. (2007). Biological information specialists for biological informatics. *Journal of Biomedical Discovery and Collaboration, 2, (1)*. Retrieved February 1, 2011, from http://www.j-biomed-discovery.com/content/2/1/1.

Hey, T., & Trefethen, A. (2003a). e-Science and its implications. *Philosophical Transactions A, 361, (1809)*. Retrieved February 1, 2011, from http://rsta.royalsocietypublishing.org/content/361/1809/1809.

Hey, T., & Trefethen, A. (2003b). The data deluge: An e-Science perspective. In F. Berman, G. Fox & A. J. G. Hey (Eds.), *Grid computing - making the global infrastructure a reality*. Chichester, UK: Wiley and Sons.

Kieserman, B. (2008). People matters. *The Bottom Line: Managing Library Finances, 21, (1)*. Retrieved February 1, 2011, from http://www.emeraldinsight.com/journals.htm?issn=0888-045x&volume=21&issue=1&articleid=1728621.

Kinkus, J. (2007). Project management skills: A literature review and content analysis of librarian position announcements. *College & Research Libraries, 68(4): 352-363*. Retrieved February 1, 2011, from http://crl.acrl.org/content/68/4/352.

Lai, L.-L. (2005). Educating knowledge professionals in library and information science schools. *Journal of Educating Media & Library Sciences*, *42, (3)*. Retrieved February 1, 2011, from http://www.doaj.org/doaj?func=abstract&id=169858&recNo=4&toc=1.

Palmer, C. L., Heidorn, P. B., Wright, D., & Cragin, M. H. (2007). Graduate curriculum for biological information specialists: A key to integration of scale in Biology. *The International Journal of Digital Curation*, *2, (2)*.

Pomerantz, J., Oh, S., Wildemuth, B. M., Hank, C., Tibbo, H., Fox, E. A., et al. (2009). Comparing curricula for digital library and digital curation education. *Proceedings of DigCCurr2009 International Symposium on Digital Curation.* Chapel Hill, NC, USA.

Promis, P. (2008). Are employers asking for the right competencies? A case for emotional intelligence. *Library Administration & Management*, *22, (1)*. Retrieved February 1, 2011, from http://journals.tdl.org/llm/article/view/1715/995.

Renear, A., Teffeau, L., Hswe, P., Dolan, M., Palmer, C., Cragin, M., et al. (2009). Extending an LIS data curation curriculum to include Humanities data. *Proceedings of DigCCurr2009 International Symposium on Digital Curation.* Chapel Hill, NC, USA.

Ricking, M., & Booth, R. (1974). *Personnel utilization in libraries: A systems approach.* Chicago: American Library Association.

Stetz, T. A., Button, S. B., & Porr, W. B. (2009). New tricks for an old dog: Visualizing job analysis results. *Public Personal Management*, *38, (1)*. Retrieved February 1, 2011, from http://www.ipma-hr.org/sites/default/files/pdf/ppm/ppmspr2009.pdf.

von Laszewski, G. (2009). Teaching about advanced cyberinfrastructure for Computational Science and e-Science. In Y. Xu (Ed.), *Transform Science: Computational Education for Scientists.* Redmond, WA: Microsoft, Corporation.

Wang, S., Liu, Y., Wilkins-Diehr, N., & Martin, S. (2009). SimpleGrid toolkit: Enabling geosciences gateways to cyberinfrastructure. *Computers & Geosciences*, *35, (12).* Retrieved February 1, 2011, from http://linkinghub.elsevier.com/retrieve/pii/S0098300409002210.

Wang, W. (2009). Computational thinking: Enabling translational research in biological sciences. In Y. Xu (Ed.), *Transform Science: Computational Education for Scientists*. Redmond, WA: Microsoft Corporation.

Waters, J., & Allen, R. B. (2009). What should we teach about METS in a digital preservation course? *Proceedings of DigCCurr2009 International Symposium on Digital Curation.* Chapel Hill, NC, USA.

Witkin, B., & Altschuld, J. (1995). Planning and conducting needs assessments: A practical guide. Thousand Oaks, CA: Sage Publications Inc.

Xu, Y. (Ed.). (2009). Transform science: Computational education for scientists. *Redmond, WA: Microsoft Corporation.*

Yakel, E., Conway, P., & Krause, M. G. (2009). Thinking like a digital curator: Creating internships in the Cognitive Apprenticeship Model. *Proceedings of DigCCurr2009 International Symposium on Digital Curation.* Chapel Hill, NC, USA.