

Jun 23rd, 1:00 PM - 2:00 PM

e-Science and data support services: a survey of ARL members

Catherine Soehner
University of Michigan, csoehner@umich.edu

Catherine Steeves
University of Guelph, csteeves@uoguelph.ca

Jennifer Ward
University of Washington, jlward1@uw.edu

Follow this and additional works at: <http://docs.lib.purdue.edu/iatul2010>

Catherine Soehner, Catherine Steeves, and Jennifer Ward, "e-Science and data support services: a survey of ARL members" (June 23, 2010). *International Association of Scientific and Technological University Libraries, 31st Annual Conference*. Paper 1.
<http://docs.lib.purdue.edu/iatul2010/conf/day3/1>

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.

E-SCIENCE AND DATA SUPPORT SERVICES: A SURVEY OF ARL MEMBERS

Catherine Soehner
University of Michigan, USA, csoehner@umich.edu
Catherine Steeves
University of Guelph, Canada, csteeves@uoguelph.ca
Jennifer Ward
University of Washington, Seattle, USA, jlward1@uw.edu

Abstract

After working for several years to promote research library support for e-science and e-research, the Association of Research Libraries E-science Working Group surveyed member libraries in the Fall of 2009 to gather data on the state of engagement with e-science issues. ARL members include the largest research libraries in the US and Canada and collectively support a substantial proportion of the North American research enterprise. The survey provides a current snapshot of data curation and e-science support activities, documenting a range of approaches and strategies evolving at research institutions and among research libraries.

The research libraries reported on widespread planning and development at their institutions and described a variety of models for service provision and infrastructure development. Library engagement in data curation was also common. Respondents provided information on library services, organizational structures, staffing patterns and staff development, and involvement in research grants, along with perspectives on pressure points for service development.

Background

As e-science has emerged as a persistent and increasingly large part of the research enterprise, academic libraries are exploring new roles, services, staffing and resources to address the issues arising as the result of this new area of research. The Association of Research Libraries recognized the need for more information and began to organize efforts to review e-science trends in detail.

The ARL E-Science Working Group was formed in 2008 as part of the recommendations from an earlier Joint Task Force on Library Support for E-Science (http://www.arl.org/bm~doc/ARL_EScience_final.pdf) that determined e-science was a continuing issue for research libraries and deserved an ongoing working group. A survey was developed by the E-science Working Group in answer to their formal charge to “build an understanding of how libraries can contribute to e-science activities in their institution” and “identify organizations and institutions that have similar interests in e-science to leverage research library interests”.

This is a preliminary report describing the results of that survey. The full report, “E-Science and Data Support Services” is available at <http://www.arl.org/rtl/plan/nrnt/esci2009rept.shtml>.

Method

The E-Science Survey was released in August 2009 to ARL member libraries as a web-based survey. Responses were accepted through November 16, 2009. Sixty-one libraries responded out of 123 total ARL member libraries. Not all 61 respondents completed the survey and not every library filling out a majority of the survey answered every question. Therefore, the percentages relayed in the Results section of this report should be seen in context of the actual number of answers to individual questions rather than as a percentage of the 61 respondents.

In addition to the survey, 7 institutions were contacted for interviews. Interviews were conducted over the phone with follow up questions sent via email. These interviews along with the corresponding institutions’ responses to the survey comprise the Case Studies section of the full report.

As with most surveys, individuals interpreted questions in a variety of ways. The topic of the survey, e-science, has a vocabulary that is still evolving. Academia as a whole is still settling on terminology and the transition will take some time, which is reflected in some of the responses to the survey and the way survey questions were interpreted.

Results

Approaches to e-science in Institutions

A high proportion of respondents to the survey indicated that their institutions were providing infrastructure or support services for e-science (45 of 58 or 77.5% of respondents). Institutional approaches to the need to support e-science vary considerably and institutions are at different stages in the development of a response to data-intensive e-science. Some universities are planning to create or already have established formalized institution-wide e-science strategies. They have formed task forces or standing committees to develop policies and central infrastructure and services to support and foster e-science and data. Some institutions still primarily see disparate departmental responses emerging, while still others have established hybrid models that provide for both institution-wide and unit specific efforts. These three institutional responses impact the nature of the collaboration and the role of the Library in e-science.

Institutional Approaches

This examination of institutional approaches to e-science will focus on those respondents who selected the following as the best description for how their institution had organized itself to advance e-science planning and policy development: "My institution has or is planning an institution-wide structure (such as a group or task force) to advance e-science planning and policy development" AND those respondents who answered: "My institution has or is planning a hybrid structure that includes both institution-wide and unit-specific efforts."

It is clear from the survey responses that institutions with or planning a predominant **institution-wide structure** to advance e-science planning and policy development were few. Only 4 respondents of 45 indicated that their institution had or was planning an institution-wide approach, such as a group or task force to advance e-science. 3 of 4 respondents planning institution-wide policy and planning bodies indicated that the groups were composed of staff from information technology, faculty researchers, the office of research, and the Library. One indicated the institution-wide planning and policy development was to be conducted by a body comprised of the CIO and the Library.

UC San Diego's Research Cyberinfrastructure Design Team (RCIDT) and their report *Blueprint for the Digital University* (<http://research.ucsd.edu/documents/rcidt/RCIDTReportFinal2009.pdf>) illustrate the strength of this approach, as well as the belief that access to a centralized cyberinfrastructure is essential to e-science and the modern research university. The UCSD Case Study in the full report provides more context for this e-science support strategy.

The majority of survey respondents with an institutional response to e-science, 27 of 45 or 60%, indicated that their institution had or was planning a **hybrid structure** that included both institution-wide and unit-specific efforts to advance e-science planning and policy developments.

Similar players were reported for the institution-wide groups developed by the institutions with hybrid strategies including staff from information technology, faculty researchers, the office of research, and the Library. The only notable difference was the rate of participation reported for the campus office of research. All other named participants were selected by 90-100% (24-26) of hybrid respondents while only 61% (16) indicated the office of research was a core member of institutional efforts.

Twenty-one of 27 hybrid institutions indicated an institutional group with e-science planning and policy responsibilities either existed or was being planned. Respondents described central groups that were either temporary in nature in the form of task groups or of a standing or permanent

nature. They also described research labs, centers and institutes with a campus e-science or cyberinfrastructure mandate.

Such institutional groups were charged with a variety of responsibilities: seeding and providing central resources in support of e-science and research cyberinfrastructure (RCI) activities; developing plans or proposals for the establishment of institution-wide RCI for e-science; and conducting interdisciplinary research into e-science and cyberinfrastructure problems. Many institutional groups take responsibility for distinct components of e-science and research infrastructure such as data management or high performance computing.

For instance, the locus for planning and services around e-science issues at the University of Washington is the UW eScience Institute (<http://escience.washington.edu/>), an interdisciplinary and institution-wide coordinating body. Library staff are involved in discussions, planning, and referral with UW eScience Institute staff regarding data curation issues. Johns Hopkins had an e-science task force that has been disbanded and can be recalled at the discretion of the Provost. Currently, the Institute for Data Intensive Engineering and Science (IDIES) (<http://idies.jhu.edu>) has become the most visible umbrella organization for e-science activities at Johns Hopkins.

The University of Minnesota's Research Cyberinfrastructure Alliance (<http://sspu-test.oit.umn.edu/rca/>) is jointly sponsored by the CIO, the VP for Research, and the University Librarian. Membership is drawn from those organizations, college information technology service organizations with robust research support, and faculty researchers. Their charge is to assess service infrastructure, policy, budget models, and opportunities to leverage collegiate infrastructure for more enterprise implementations.

At the University of Utah two institutional groups exist: the Cyberinfrastructure Council and the Knowledge Management Committee (<http://www.it.utah.edu/leadership/committees/Cyber/index.html>). The Council is involved in high performance computing, data centers, other computing and network issues. The Knowledge Management Committee is more oriented to the content of e-science and data curation, leveraging the intersection with the institutional repository and scholarly communications initiatives. Data curation is integral to data center operations and so in this arena both the Council and the Knowledge Management Committee share responsibility.

Unit-by-Unit Approach

Virtually all research institutions have large grant-funded projects that have developed their own infrastructure on a college, departmental or unit level. The respondents who presented institutional collaboration models are differentiated in the section above because they indicated the existence of institution-wide or central activity and collaboration in the planning and policy development for e-science at their institution. This section, Unit-by-Unit Approach, focuses on the respondents that selected the following as the best description for how their institution had organized itself to advance e-science planning and policy development: "At my institution individual units (i.e. departments, colleges, schools, etc.) develop infrastructure and policies related to their own e-science needs."

There were a total of 12 survey respondents (out of 45) who self identified their institutions as having individual units developing infrastructure and policies related to their own e-science needs. These respondents describe their organizational culture as being highly decentralized with the focus of e-science activities in particular subject areas or interdisciplinary institutes. While some survey respondents certainly indicated a strong desire on the part of libraries to participate more fully in e-science efforts on campus, only 5 of those 12 indicated that their library was providing infrastructure or support services for e-science. The remaining 7 said that "e-science infrastructure or support services are in the planning stages." This pattern, while based on a small number of respondents, may indicate one of the difficulties libraries experience when attempting to collaborate with other campus units to develop a centralized service point on a campus that has a decentralized culture.

Multi-institutional Grants and Collaboration

While the survey largely focused on institutional activities, there were several indicators in the survey responses pointing to multi-institutional responses. Cooperation between institutions, including between libraries of different institutions, is becoming increasingly necessary and is likely to become a practical method of addressing common issues in e-science support. Approximately half of the respondents (20 of 41) indicated that their institutions were involved in a collaborative program with another institution in support of e-science. Library involvement in these collaborations was not guaranteed, but was frequent among those filling out the survey. Of those 20 respondents with collaborative programs, 16 indicated that the library was involved in some way.

When describing their multi-institutional collaborations, some respondents described their connections with other institutions as a result of joint grant proposals, many of which were NSF DataNet grants. There were equal numbers of respondents (19 out of 42, or 45%) whose institutions were involved in a DataNet proposal and those that were not. When providing more details about their DataNet proposals, 11 institutions indicated that more than one institution was involved. When asked if the library was involved in the DataNet proposal 56% (17 of 30) stated that the library was involved in the proposal. For example, the University of California at San Diego listed their "Datapedia of Science" grant proposal as a multi-institutional project that will provide "an innovative platform for the long-term, scholarly publication and preservation of scientific data." Purdue University made one additional comment that was significant: "The relationships developed in preparing this proposal were very beneficial in increasing the visibility of the Libraries on campus and its potential involvement with collaborating on data management issues. The credibility and involvement, especially within science and engineering, increased in many areas." This theme of relationship building between librarians and researchers is highlighted throughout this report as an activity that was most productive in developing future partnerships and collaborations.

Library participation in grant proposals outside of the NSF DataNet grants was also prevalent and included such interesting proposals as the one from UIUC, "Digging into data", a grant focusing on earthquake engineering, and another proposal from the University of Oregon, in which molecular biology faculty requested funding for a bioinformatics center for genomics research that will include space for a librarian. Another form of collaboration highlighted in the survey was library to iSchool collaborations, some of which involved more than one institution. Many respondents made reference to the summer institute on data curation offered by UIUC taking into account partnerships to develop the institute or further develop internships related to the institute. Other respondents described cooperation outside of grant proposals that provided some connection around e-science across institutions. The Texas Digital Library and the California Digital Library were cited as multi-institutional collaborations that involved infrastructure in support of e-science. Other efforts outside of grant proposals include organized discussions about e-science issues among several institutions.

Data Support and Services in Institutions

Survey respondents with institution-wide, hybrid and unit specific e-science planning structures were also asked questions specific to data support and services developments at their institutions including data curation, data needs and resource assessments, data center(s) and digital lab notebook applications.

The majority of respondents to the questions pertaining to data curation and support (total of 42 responded to these questions) indicated that there were no designated units to provide data curation and support for scientific research data on their campus but as many as 19 (45%) of these institutions did indicate the presence of designated unit(s). The nature of the units identified included: data centers, disciplinary informatics centers and institutes, statistical analysis and academic computing centers and services, library data and institutional repositories, digital research and curation centers, campus information technology units, and, high-performance computing and cyberinfrastructure centers and institutes.

Twenty-two respondents indicated that they had both central and distributed data centers for research data on their campuses, while 18 institutions purport to have a distributed data centers system. This reflects the distributed nature and culture of decentralization in large research institutions. A single respondent indicated that they had a central institutional data center. The comments indicated that some institutions are striking task forces to examine data center consolidation. Others have established institutional data centers for e-scientists and data intensive researchers and “encourage” them to avail themselves of their services. The survey responses did not clarify what incentives or policies were being used to provide such encouragement and enable transition to central research data center use.

The *Research Data Management and Publishing Support at Cornell* website provides an excellent window onto both the complexity of data management on campuses and the resulting importance of identifying and designating units who support data management and curation for researchers (<https://confluence.cornell.edu/display/datasupp/Home;jsessionid=B99704EAF8C509A4C1D2D4D7F7E75B8E>).

At the time of the survey, 62% of respondents (26 of 42) indicated that their institution had not conducted an assessment of data resources and needs. Many however were planning assessment activities while 16 respondents had conducted assessment activities. For example, UCSD conducted a needs assessment for research data management, analysis, storage, archival, and access as a part of their 2008 Research Cyberinfrastructure Survey (Appendix D of the aforementioned report *Blueprint for the Digital University* <http://research.ucsd.edu/documents/rcidt/RCIDTReportFinal2009.pdf>). Another example can be found in the University of Wisconsin’s *Report of the Research Data Management Study Group* (<http://minds.wisconsin.edu/handle/1793/34859>).

The survey also explored institutional digital lab notebook application support. Only 5 institutions (12%) indicated that their institution supported them (Texas A&M University, Canada Institute for Scientific and Technical Information, Yale University, McGill University, and Purdue University). Support was provided by departmental IT, teaching support services, and most often, by individual researchers who self-support the use in their labs. Twenty eight percent indicated that their campus did not support digital lab notebook applications. However, the majority of respondents to this question, 60%, were not sure whether such support existed at their institution, suggesting a low level of awareness of these activities amongst research libraries. None of the respondents knew of archival services for digital lab notebook applications.

Approaches to E-science in Research Libraries

Approximately 73% (29 out of 40) of the respondents indicated the library was involved in e-science support at their institutions and that leadership of these efforts was primarily through a team effort (15 out of 31 respondents) or some combination of individuals, units and teams working together (13 of 31).

The naming of these groups/teams/committees ranged from the typical e-science or e-Research names to a focus on data curation to names that were clearly a reflection of a particular project or department. The titles of the position designated to lead these teams were wide ranging. Almost half of the titles listed were at the Associate Dean level, with the next most popular title being a science librarian. Most groups were formed and began their work in 2008 or 2009, with one having begun as early as 2006.

When asked to describe the library organization for developing e-science plans and programs, only 2 respondents indicated that there was a specific library unit focused on e-science issues. Emory University has the Digital Programs and Services unit within their library and Johns Hopkins University Library named the Digital Research and Curation Center as their primary unit within the library that provides e-science support (<http://ldp.library.jhu.edu/dkc>). Beyond those two institutions, there were a wide range of responses with different levels of formality and a variety of stages of development. For example, on the side of formal and well-developed programs, one respondent

described their e-science efforts as follows: “The Associate Director for Technology works on the long-term strategy for science data curation, including assessment of current needs and appropriate role for the library, as well as the technological infrastructure required; there is a public services committee who are developing expertise in the topic, talking to faculty, developing pilot archiving projects, and teaching one-hour courses on the subject to students. There are[sic] also a digital product manager and metadata expert assisting all of this. Overall, lots of people are involved in some way.” On the other side of the spectrum were statements such as, “We are very early in our planning...” and “Informal, evolving structure...” and “Planning in progress to develop data services positions.”

Approximately 87% (27 out of 31) of the e-science services offered by libraries are provided in collaboration with other units on campus. Of the 26 respondents who provided more detailed information regarding campus units, half cited a working relationship with the campus or centralized IT organization. Six libraries indicated working with the Vice Provost for Research (or similarly named unit) and 8 libraries reported working with a variety of individual departments. When asked about the subject disciplines of the individual campus departments, the usual suspects emerged such as biochemistry/molecular biology, biomedical engineering, chemistry, computer science, environmental science, earth science and health. The more unusual subject areas included management, education, Latin American studies and biological anthropology.

A good example of these kinds of library-campus collaborations can be seen at the University of Oregon where the Library’s Metadata Services and Digital Projects (<http://libweb.uoregon.edu/catdept/home/>) is combining efforts with the Campus Information Services to make an “Inventory of science data sets across these departments: biochemistry/molecular biology, biology, biological anthropology, chemistry, computer and information sciences, environmental science, geography, geology, human physiology, physics, and psychology.” Another example at the University of British Columbia combines the efforts of the Library’s Institutional Repository and Scholarly Communications with the Office of Research Services. These two groups “are at the beginning stages of exploring how to handle data associated with research. In particular, they are considering the mandate of the Canadian Institutes of Health Research.”

Reference and Consultation Services

Consultation services such as identifying data sets, providing access to data and articulating current standards for organization of data in specific subject areas seem to be a natural fit for subject librarians who provide similar services for other types of information. This perception was confirmed in the survey results where a clear majority of respondents provided some combination of the following services for researchers on campus:

- Finding and using available technology infrastructure and tools (22 of 29 respondents)
- Finding relevant data (24 of 29)
- Developing data management plans (23 of 29)
- Developing tools to assist researchers (22 of 29)

A few described more advanced services such as “archiving relevant data and curating it for long-term preservation and integration across datasets” and “providing curatorial and data stewardship services” as part of data management plans.

Several libraries listed raising awareness as another key activity and have created websites dedicated to describing the e-science services they provide, for example, Cornell University Library’s list of data management services on campus: <https://confluence.cornell.edu/display/datasupp/> and the University of Oregon Libraries description of data services: <http://libweb.uoregon.edu/faculty/SciDataInfo.html>

Only a few libraries (8 out of 30 respondents) provided workshops for faculty regarding e-science issues and several of those were in the planning stages. A couple of the workshops focused on the

use of data in Geographical Information Systems while others cited data management, tools and best practices. However, when the survey asked, "Does your library include policy issues associated with e-science (e.g., open data, compliance with federal agency policies) in its outreach program?" many more answers (17 of 31) revealed connections with faculty regarding open access/open data, NIH compliance as a part of their scholarly communication efforts. Most libraries are very familiar with copyright and intellectual property rights questions coming to them and even offering workshops on these topics. The libraries may not see these scholarly communication issues as being connected to e-science, when, in fact, the connection is closer than is realized.

In addition to consultation and reference services, many libraries (20 out of 31 respondents) manage or participate in managing technology related to e-science such as servers for data storage and tools/software for analysis. When asked to provide details, most libraries reported the provision of servers for data storage and many reported this storage capacity in context of a specific project, including DISCOVER (<http://arecibo.tc.cornell.edu/DRSG/Default.aspx>), VIVO, DataStaR (<http://datastar.mannlib.cornell.edu/>), Chronopolis (<http://chronopolis.sdsc.edu/>), Harvest Choice (www.harvestchoice.org), EthicShare (www.ethicshare.org) and DataONE (<http://dataone.org/>).

Resourcing E-Science Activities in Libraries

When asked, "Who provides/will provide reference/consultation services to researchers?" more than half (17 out of 29 respondents) indicated that they have both individual discipline librarians/staff and dedicated data librarians/specialists taking on these duties. Libraries reported using a mix of strategies to create a workforce with the skills and capacity to provide e-science services and programs. Respondents were asked to identify all strategies that applied to their institution's approach. Considering the current economic climate, it is not surprising that a majority of libraries (18 out of 28 respondents) are reassigning existing staff or providing training to existing staff as part of an overall strategy to incorporate e-science responsibilities into their current portfolios. In addition to reassigning existing staff, libraries have hired or plan to hire staff specifically to provide e-science services as part of an overall strategy. Libraries have traditionally used this combination of hiring new staff and re-training current staff to take on new areas of responsibility. This investment of resources even during budget cuts indicates the level of commitment to e-science services by many of the respondents to the survey.

Survey respondents were asked to provide details for three positions that have or will have data management or e-science-type duties as a major part of their portfolio. A total of 65 positions were described, most of which were permanent positions (only four were grant funded or temporary positions). The most popular titles for these positions included the word "data" (31%), while the next most popular title was a subject specialist (20%) closely followed by managers or directors of digital repositories (17%).

The question of the value of the MLS degree has been debated, particularly as technology continues to rapidly change the library world. Among the 65 positions that were reported, 64 of them provided educational degrees for the survey. Of those, 46, or 72%, indicated degrees in Library and Information Science at the Masters or PhD level for positions that were current or planned. Some of these MLS degrees, 6 in total, were paired with a discipline Masters, one was paired with a PhD and one suggested some combination of an MLS, discipline Masters, and discipline PhD. Two library science degrees were at the PhD level or reflected enrollment in a library science PhD program.

A vast majority of respondents (28 out of 31 respondents) indicated that library staff were given opportunities to develop skills related to e-science. Of these 28 respondents, only 6 indicated that they provided the opportunity to take coursework related to e-science or data management in a discipline. This small response may be a result of many institutions not having or not being close to an iSchool or not providing relevant coursework through other schools or institutes. However, if appropriate coursework can be identified, library administrations should consider making these opportunities available to staff as a method of retraining or continuing education.

Pressure Points

The top three areas identified by survey respondents as pressure points include a lack of resources, difficulty acquiring the appropriate staff and expertise to provide e-science and data management or curation services, and the lack of a unifying direction on campus. Although not mentioned as frequently, the lack of infrastructure to handle, preserve and provide access to data was another area of stress as libraries consider their e-science roles on campus. A somewhat surprising answer from seven different respondents was a lack of faculty interest in data issues as a major source of pressure. Before libraries can play a credible role in e-science and provide data management, curation and preservation services, there must be an identified need by the campus. Continued connection with faculty about other library services that faculty see as relevant will provide an avenue for discussions and education around issues regarding data curation, preservation and access.

The relative recent arrival of e-science support through libraries and a desire for more expertise in this area begged the question of information exchange between ARL libraries. A majority of respondents, 48 out of 53, indicated that they were willing to participate in an information exchange, but only 18 of those felt they had enough experience with e-science support to have something worthwhile to offer. Topics of interest were primarily around best practices such as staffing levels, descriptions of projects and services, policies, successes, grant opportunities, and how libraries established expertise on campus.

Observations

Collaborations are essential to address even modest support of e-science. The survey revealed successful and frequent collaborations on all levels: between libraries of different institutions, between libraries and the departments they serve, between various departments to address interdisciplinary subject areas and between institutions. Because the data sets created by modern scientific methods are often very large, the resources required to manage that data must also be extensive. For this reason, collaborations will continue to be an important method to address the enormity of the challenges posed by e-science.

Faculty interest and institutional support at the administrative level are important for success of library services in this area. Without faculty and institutional agreement, libraries will find themselves preaching about the importance of data curation, preservation and access without making an impact. Our institutional repositories and the lack of faculty contributions to them are good examples of this phenomenon. We can preach about the benefits of our IR's but without faculty or institutional acceptance, there will be limited use. All of the case studies described in this report demonstrate this point. University of California San Diego, Cornell, Johns Hopkins, and Purdue University Libraries are involved in data management activities at campuses where the administration of the campus as well as some faculty (although not all) recognize the importance of data management.

The Masters of Library and Information Science degree has a place in this new area of librarianship. Of 64 positions described that provided degree requirements, 46, or 72%, listed Library and information Science at the Masters or PhD level that were either in place in current positions or were planned for future positions. Since reassigning staff is a major strategy for resourcing roles in e-science in libraries, it is not too surprising that the MLS has shown up in the survey responses as a degree found in a majority of data positions.

There is some evidence in recent job postings that the focus on the MLS is in flux. As can be seen in Appendix II, one recent posting required an "ALA accredited master's degree in library or information science" while another required "Demonstrated expertise in data management or information science. This would preferably take the form of direct experience with data curation/management, but could include an M.L.S./M.L.I.S. degree with an emphasis on data management." Interviews with Case Study participants highlighted success of librarians in e-science services with advanced degrees in science or engineering. Connections between faculty and librarians appeared easier to begin and sustain if the advanced science or engineering degree

had been obtained from the institution in which the librarian was now employed. At the same time, the cooperative and team oriented values of most MLS degree holders along with the understanding of the role and importance of metadata and preservation makes the combination of the MLS along with an advanced science degree a useful and successful combination.

The fact that investments in e-science activities are being made even during difficult budget times demonstrates that this is a priority for libraries. The priority in libraries is to remain relevant to the institutions they serve. As research activities become more data intensive and as faculty and institutions become increasingly concerned about the preservation and access to that data, libraries have an opportunity to demonstrate their expertise and relevance to their institutions and should take advantage of opportunities to move forward into data management activities. There is a parallel between information resources management (our traditional role) and data management. Collecting and preserving information for researchers is a recognized role of libraries and, therefore, libraries are an obvious partner to provide the expertise to assist the university with a centralized plan for data management.

There is some concern that if libraries do not act quickly, others (publishers or vendors) will collect the data and then charge universities a lot of money to get it back. Others are concerned that a lack of quick action by libraries will lead to their becoming irrelevant much faster than previously imagined. Research libraries are positioned to step up to the opportunities to move forward as e-science becomes a usual part of research practices and as faculty and universities recognize the need to create structures to curate, preserve and provide access to the results of that research.

Strategies for data curation, management and preservation are still young and evolving. DataNet grants through the National Science Foundation as well as other externally and internally funded programs will provide substantial models and information that will help guide decisions over the next few years.

Conclusion

The results of this survey indicate that engagement by libraries in the areas of e-science have been developing rapidly in the past few years. This has ranged from answering basic questions about metadata and open access standards to providing infrastructure for curating and managing large datasets. Collectively among the respondents there is a more sophisticated view of data management skills, services and recourses, and promising strategies for engagement in these activities are becoming clear and are often highly collaborative. The investment of resources in e-science even during difficult budget times indicates a strong priority among libraries and institutions. However, the size of the issues involved demands collaboration between institutions and libraries to solve the collective data problems. As stated by Susan Parham of Georgia Tech, "This area is very important, but is much larger than a single institution. We need a national framework for addressing the management, re-use and preservation of scientific data."

Bibliography

- Association of Research Libraries Joint Task Force on Library Support for E-Science. (2007). *Agenda for developing e-science in research libraries: Final report and recommendations to the scholarly communication steering committee, the public policies affecting research libraries steering committee, and the research, teaching, and learning steering committee.* Retrieved from http://www.arl.org/bm~doc/ARL_EScience_final.pdf.
- Association of Research Libraries. (2006). *To stand the test of time: Long-term stewardship of digital data sets in science and engineering.* Retrieved from <http://www.arl.org/bm~doc/digdatarpt.pdf>.
- Association of Research Libraries. (2007). *Agenda for developing e-science in research libraries: ARL joint task force on library support for e-science final report & recommendations.* Retrieved from http://www.arl.org/bm~doc/ARL_EScience_final.pdf.

- Atkins, D. E., Droegemeier, K. K., Feldman, S. I., Garcia-Molina, H., Klein, M. L., Messerschmitt, D. G.,...Wright, M. H. (2003). Revolutionizing science and engineering through cyberinfrastructure: report of the national science foundation blue-ribbon advisory panel on cyberinfrastructure. Retrieved from http://www.nsf.gov/publications/pub_summ.jsp?ods_key=cise051203.
- Gastinger A. (2009). Upgrading the eLibrary: Enhanced information services driven by technology and economics. *Library Hi Tech News*, 26(1-2), 1-5.
- Hey, T., Tansley, S., Tolle, K. (Eds.). (2009). *The fourth paradigm: Data-intensive scientific discovery*. Retrieved from <http://research.microsoft.com/en-us/collaboration/fourthparadigm/contents.aspx>.
- Hey, T., Hey, J. (2006). E-Science and its implications for the library community. *Library Hi Tech News*, 24(2), 515-528.
- The Fourth Paradigm: Data Intensive Scientific Discovery", has been released by Microsoft Research and is available for download at <http://research.microsoft.com/en-us/collaboration/fourthparadigm/contents.aspx>.
- Lynch, C. A. (2003). Institutional repositories: Essential infrastructure for scholarship in the digital age [Electronic Version]. *Association of Research Libraries*, 226, 1-7.
- Rambo, N. (2009). E-Science and biomedical libraries. *Journal of the Library Medical Association*, 97(3). doi: 10.3163/1536-5050.97.3.001.