Academic Library, e-Science/e-Research, and Data Services in a Broader Context

Academic Library, e-Science/e-Research, and Data Services in a Broader Context

Minglu Wang

In North America, academic libraries' data services have recently emerged as a new field during a very interesting time when academic libraries are adjusting themselves to be able to deal with more digital material and at the same time trying to be further involved in the academic research process. The academic world is also undergoing a transformation into a new paradigm of doing research called e-science, which is characterized by data-intensive and networked research. Managing and curating the ever-increasing amount of data seems to be a natural extension of the established function of libraries. However, if we look at the broader landscape of e-science and all the supporting systems that are under development accordingly, we will see more than one model of data services provided by different institutions, or by different combinations of institutions. This paper will summarize these two major e-science support models, both in North America and in Europe. For both data services planners and practitioners, we need more contextual learning about the academic world and the emergence of the e-science paradigm, and a more visionary view of libraries among all the services departments/agencies around us so that we all can better design our academic library services and continue to promote and develop it.

Introduction

The academic world is currently undergoing a period of transformation, involving the new research paradigm of e-science and e-research, and characterized by data-intensive and networked collaboration. At the same time, academic libraries are adjusting themselves to dealing with more digital material and trying to be further involved in the academic research process. This presents a unique opportunity for academic libraries to reinvent themselves based on the new e-science/e-research needs and issues that are generated in the research communities that we serve.

Over the last few years, new data service programs have emerged rapidly in academic libraries in the United States, which have worked hard to support data access and management, and which have approached institutional data curation as a natural extension of their traditional information organization and dissemination functions. But academic libraries are not always active initiators of data curation, especially in the current European context, in which national data centers have been developed to take on the responsibility of preserving publically-funded research data. However, even in such cases, academic libraries have the potential to act as effective mediators between researchers and information and communication technologies, funding agencies, and different entities within the centralized e-science/e-research infrastructure. Based on literature and reports that discuss current e-science/e-research development

Minglu Wang is Data Services Librarian in the John Cotton Dana Library at Rutgers, the State University of New Jersey; Email: minglu@rutgers.edu

and academic libraries' visions for our future, this article assesses trends in e-science/e-research and other aspects of new scholarly behavior that our academic libraries are particularly well-equipped to participate in and support, even beyond curating the final research products.

A Decade of E-Science/E-Research Development: National Funding Agencies' Initiatives

Although we often talk about e-science/e-research as a new academic practice or way of doing research, as a social phenomenon, it is actually inseparable from a series of past national research funding programs and initiatives, especially in the UK and US. A review of the brief history of nationally-led e-science/e-research development reveals much about the origin, current status, and future potential of the field. Importantly, not only has the terminology moved away from a focus on e-science to include e-research, but the structure and focus of national investments have also been adjusted, with projections changing accordingly.

Origins and the Trend of Extension to All Disciplines

E-science, as a term describing the increasing amount of large-scale science being conducted through global collaboration and the use of the Internet, was first introduced when the UK Office of Science and Technology launched a funding initiative in 1999.[1] In 2001, the UK National e-Science Center (NeSC) was established, and individual subject research councils began funding e-science projects within their own fields. In the US, the National Science Foundation (NSF) published a report in 2003 that used the term "cyberinfrastructure" to characterize the new information and communication technology that was supporting revolutions in the fields of science and engineering.[2] Since 2005, government approaches to cyberinfrastructure have moved beyond science and engineering and toward a NSF-wide cross-directory grant program, but have remained very science-centric in terms of the projects that have been supported.[3]

While e-science and cyberinfrastructure were initially limited to describing hard science involving large-scale data that required high-capacity computing power to process, e-research was developed as a term that includes research practices in social sciences and humanities. These fields have been no less impacted by new information and communication technologies, and are also characterized by networked and data-intensive features. In *e-Research: Transformation in Scholarly Practice*, Jankowski generally defines the major dimensions of e-research as "computerization, virtual organization structure, internet-based tools, visualization, and the publication, distribution, and preservation of scholarship via the Internet".[4]

The change in terminological focus from e-science to e-research demonstrates one major trend in the support system for new academic research practice, in its expansion beyond the fields of science and engineering and into the social sciences and humanities. In 2005, the UK's Arts and Humanities Research Board became the Arts and Humanities Council (AHRC), and joined the other UK research councils in supporting general e-science programs, with a focus on e-humanities.[5] In 2006, the US National Endowment for the Humanities (NEH) launched the Digital Humanities Initiative, which in 2008 became a designated office with its own budget line.[6] In terms of e-social science, in 2004, the UK Economic and Social Research Council (ESRC) established the National Center for e-Social Science (NCeSS), as a part of the country's e-science program.[7] While in the US, e-social sciences has not yet earned widespread attention as an independent emerging field, in 2012, the NSF's Cyberinfrastructure program began designating grants to support research in the fields of social and behavioral sciences.[8]

Different Models of E-Science/E-Research Support and Challenges for Future Development

Although both the UK and US are actively developing national funding programs to support e-science/e-research infrastructure and tools, they have otherwise adopted different strategies. With efforts led by national e-science and e-social science centers and research councils on all subjects, the UK has followed a clear national vision and design from the very beginning. In the US, the NSF initially encouraged bottom-up community engagement in the hope that infrastructure would aggregate upward like a pyramid, but later determined that community-level practices were too complicated, and that a spiral model involving a system gradually increasing in sophistication was more realistic.[9] However different the models are in the UK and US, they face a common challenge: the wider range of adoption of developed systems and tools among researchers.

In a 2009 report reviewing three phases of an e-science program, the UK's research councils expressed their sense of accomplishment, in terms of the strong foundation that had been laid for future transformative research.[10] Not only is national-level infrastructure now in better shape in the country, but institutional-level nodes and network infrastructure are well distributed and connected. Future challenges include continuing to fund the existing system and preparing for even larger volumes of data, and addressing the fact that individual researchers are still less connected with each other and with the infrastructure and tools that have already been developed.

In the US, although the community-based development strategy has attracted active participants from many research disciplines to work on e-science/e-research projects, tools, and infrastructure developments, collaboration is still limited, national-level coordinated infrastructure is very weak, and as in the UK, researchers lack incentives to adopt new tools and software. In 2009, the NSF Advisory Committee on Cyberinfrastructure (ACCI) stressed the importance of the higher education community and emphasized the need for members to participate in a coherent cyberinfrastructure and support and reward researchers' participation.[11]

Data Curation and Management Support Strategies: the UK And US

Although data curation and management are natural and central components of e-science/e-research development, institutions and agents in the UK and US are reacting to the challenges and opportunities involved in different ways. In the UK, national data centers are more active in developing data curation and data management infrastructure and services, while in the US, academic libraries are on the front lines, investigating the data service needs of researchers, developing support systems to assist them, and addressing their data curation and management needs.

UK National Data Centers, Centralized Data Curation, and Management Support

In the UK, research councils in both the fields of science and social sciences are the major funding agencies for a group of well-established national data centers, including the UK Data Archive, UK Solar System Data Center, and Natural Geoscience Data center.[12] In addition to helping to develop national-level research data repositories, these research councils also create data management policy regulations that enforce data sharing, and which have become increasingly explicit in addressing the roles, responsibilities, and costs and benefits involved.[13] Furthermore, the UK research councils, together with higher education funding bodies, collectively fund data management support projects that can be directly used in UK research institutions. JISC (historically Joint Information Systems Committee, now a company) has been a centralized gateway for such funded projects. JISC has funded many aspects of data management support projects, including the development of training materials and data management infrastructure and data management planning tools.[14] According to the recent announcement of a funded program that includes a data registry supporting data discovery and reuse, UK research library societies and associations have also become major stakeholders.[15]

In 2010, Research Libraries UK (RLUK) conducted research on evolving research needs, and on the roles and skills that subject librarians require to meet these needs, based on surveys of a group of libraries with a good international coverage.[16] The report identifies nine key areas, of which six are related to data and metadata. Subject librarians tend to lack knowledge and skills in these areas, and this should be considered in future training and recruitment activities. The report encourages subject librarians to go "into a world beyond information discovery and management, collection development and information literacy training, to one in which they play a much greater part in the research process and in particular in the management, curation and preservation of research data, and in scholarly communication and the effective dissemination of research outputs."

US Academic Libraries' Data Curation and Management Support

Within the bigger framework of e-science/e-research development, the US academic library community has been very consciously observing trends, bravely envisioning new support service areas, and practically experimenting with new librarianships through the data librarian and e-science librarian positions that have been rapidly established in many libraries. In 2006, the US Association of Research Libraries (ARL) appointed a task force to address e-science issues under the umbrella of the Steering Committees for Scholarly Com-

munication and for Research, Teaching, and Learning. The e-science task force gave its final report in 2007, making recommendations for the support of e-science in research libraries, systematically introducing both national and international contexts of e-science, and encouraging academic libraries to engage in all three aspects of e-science development: data issues, virtual organizations, and policy development areas.[17] ARL maintains an e-science/e-research web portal collecting relevant articles and other resources, and also organizes e-science institutes that educate academic library leaders and familiarize them with relevant topics.[18]

In 2009, ARL conducted a survey of its member libraries in the US and Canada, concerning their participation in their campuses' e-science/e-research services. The report showed that academic libraries have only gradually started becoming active partners with institutions. While e-science/e-research support groups, data access, curation, and preservation are areas that library services can naturally fit into, libraries have not been very effective in influencing related policy developments.[19]

In 2012, the Association of College and Research Libraries (ACRL) conducted a survey of its members in the US and Canada, concerning their current states and future plans for research data services.[20] The resulting report demonstrated a strong connection between academic libraries' traditional practices of helping users locate resources and the currently most common data services: web guides that help locate data. Correspondingly, it found that most libraries were reassigning their existing librarians to new data service tasks. The report also found that only larger or doctoral-granting institutions were likely to offer additional informational/consultative and technical/hands-on research data services. It identified the data management requirements of research funding agencies (e.g. NSF) as the driving force for new research data services, and urged academic libraries to act on this opportunity quickly, keeping in mind that other campus units could reorganize and compete to provide similar services.

Academic Libraries and E-Science/E-Research Support: What's Next?

The above summaries of research reports on current academic libraries' data service support practices show that visionary library leaders from both the UK and US are advocating for academic librarians to take on more active and participatory roles in new e-science/e-research processes, in order to assist researchers in working more efficiently. For example, Medha Devare prescribes that librarians could be "middleware agents between systems and systems, and between people and people."[21] Despite this, there are still big gaps in librarians' levels of knowledge and skills, e-science/e-research services are limited to data curation and management support, and new programs are being developed only by a few bigger research libraries.

The UK's data management support system demonstrates that national data centers supported by national research councils and national digital preservation organizations (e.g. JISC) can also be active and effective agents that develop infrastructure and coordinate training for researchers, and that academic libraries can be reactive and collaborative partners in the process. However, even the UK model makes it clear that when a national initiative and visionary effort needs to practically influence researchers in their institutions and communities, academic libraries are a necessary connecting point between national regulations and requirements and individual practitioners. Academic libraries systematically reside in the vital positions between researchers and their broader research environments, which means that if librarians can be more sensitive, conscious, and responsive to the bigger trends of e-science/e-research development, we will have greater opportunities to thrive in the future by developing broader and deeper new relationships with the researchers that we serve.

Opportunities for Academic Libraries, Based on Structural Position and Historical Experience

The gaps and disconnections that exist between national e-science/e-research infrastructures, existing tools, individual researchers, and everyday practices have been identified by both the UK and US national funding agencies as major challenges for future development.[22] Current discussions have focused on providing researchers with both "carrots" (e.g. tenure promotion recognition) and "sticks" (e.g. funding agency policies and regulations), encouraging them to adopt new e-science/e-research best practices. However, researchers do not mechanically follow directives based on carrots and sticks, and need to be understood, supported, and possibly educated on facing new challenges in their work. Scientists also need partners and

collaborators with whom to collectively pursue their research interests as they explore, discover, and serve the public. No existing agents are in a more perfect position that academic librarians to make connections that facilitate the technology adoption process.

Academic librarians have already developed very close relationships with the researchers in their institutions through traditional reference services and bibliographic instruction services. Academic libraries are also experienced in serving diverse campus-wide user groups, and many are partners in regional consortiums and collaboration systems that involve multidisciplinary collaboration. Most academic libraries now combine their traditional technical services and public services with new digital library programs. In employing their new digital material preservation expertise, librarians are also speaking technical and research subject languages, and designing and implementing digital asset management systems. These skills could easily be adapted to help researchers manage new materials and projects, assisting them in identifying and adopting new e-science/e-research infrastructures and tools.

One good example of the new partnership role that librarians can take on is demonstrated by a report from Purdue Librarians, which indicates that from 2005 to 2006, 11 of the university's librarians "participated in nine multidisciplinary proposals, including National Institutes of Health (NIH), National Science Foundation, and locally funded grants. It appears likely that more library science expertise can be leveraged to meet cross-multi-interdisciplinary research needs."[23] The Purdue Libraries also have a data curation center led by librarians with information system development experience, who are good at connecting technology with user needs.

Another good example of the type of connection that librarians could make, which goes beyond data curation and into everyday data collection practices, has been presented by Kristin Briney, in a talk on how librarians can help researchers identify and adopt ELNs (electronic lab notebooks).[24] Librarians can help researchers identify the important features of ELNs that support research data management best practices, and train them to implement those best practices and develop proper ELN skills when they first start to use ELN. They could also collect updated information and resources related to good ELN options that fit researchers' specific research area needs.

Supporting E-Science/E-Research Beyond Scientific Data and Curation Management

The history of social science librarians' support for electronic government data access and discovery, which dates back to the 1990s, makes it clear that academic librarians' data services are not at all new.[25] Traditional data services were provided mainly by social science subject librarians, based on the reference philosophy that is part of traditional library services theory. In contrast, data curation and management services are new areas in academic libraries, especially in terms of the involvement of new digital library departments and traditional system librarians' expertise that go beyond bibliographic services related to references and subjects. Academic libraries may now be able to move past their traditional identities as stewards of the final products of research such as publications and data, and into the complete research life cycle, participating in more aspects of academic research. Other aspects of e-science/e-research development have not yet been widely analyzed in academic libraries, and could further broaden our views on research practices that are relevant to academic libraries, and thus potentially help us identify new service and partnership areas.

Although e-science/e-research has always involved at least two major aspects: data and collaborative virtual research environments (VREs), discussion about e-science/e-research among academic libraries in the US has primarily focused on data aspects, and particularly data curation and data management plans. Only a few visionary suggestions have been made so far, for example concerning librarians' possible roles in VREs. For example, the British scholar Alan Masson has pointed out that librarians have been actively supporting and adding value to virtual learning environments (VLEs) and online class management systems, and thus could establish closer working relationships with researchers in VREs as well, developing information discovery and management tools and services that could be integrated into the workflow of the research communities.[26]

Despite the popularization of the more general term "e-research" as a replacement for the more limited "e-science," in academic libraries, new data curation and management support programs continue to focus on science rather than social sciences and humanities. This is partly due to the existence of a relatively mature social sciences data archive in US, namely the

ICPSR (Interuniversity Consortium for Political and Social Research), and partly because of science researchers' practical need to develop data management plans that are more strictly required by NSF. It is thus important to consider how academic librarians could more actively participate and assist researchers in the social sciences and digital humanities. Social sciences and digital humanities data are not hard to manage in terms of volume, in comparison to scientific data. However, social scientists and humanities researchers try to employ innovative approaches to presenting and visualizing their data, communicating their research results with each other and the public, with the help of new technologies developed in the larger e-science/e-research context. We have not seen many examples of new academic library services developed in these areas, with the exception of some new positions, such as digital humanities librarians and GIS specialists within forward-looking academic libraries. Academic librarians are in an ideal position to connect social sciences and humanities researchers and open up new possibilities for conducting and communicating research, if we are brave enough to grasp our opportunities to engage with the researchers that we serve, bringing them out from their small academic circles and into the evolving context and environment of new research.

Conclusion

Academic libraries have many means available to reinvent our programs and services based on our newly identified roles within research institutions. We can discover new research products that expand on existing structures and services and apply our traditional expertise in areas such as curating books and publications to new research materials such as research data. However, in doing so, we might also need to think outside of our traditional identities as curators of materials, and envision ourselves as active educators on new information and data literacy, acting as bridges between researchers and fast-changing research infrastructures and technologies, and as research partners in charge of research information management. With this in mind, academic libraries need to pay more attention to every aspect of e-science/e-research developments, including those that are historical and comparative, so that we understand our positions and importance within a broad holistic context. In addition to new program development, our research agendas should extend beyond practical concerns and into broader studies of the research environment, including its historical and sociological contexts and organizational and behavioral mechanisms, so that we can continuously adjust ourselves and stay relevant.

Notes

1. John Taylor, "Defining e-Science," accessed February 14, 2013, http://www.nesc.ac.uk/nesc/define.html.
2. Daniel E. Atkins et al., *Revolutionizing Science and Engineering through Cyberinfrastructure: Report of the National Science Foundation Blue-ribbon Advisory Panel on Cyberinfrastructure*, (2003), accessed March 23, 2013. http://www.nsf.gov/od/oci/reports/atkins.pdf.
3. *FY 2009 Budget Request to Congress: Office of Cyberinfrastructure*, 2009, accdessed March 23, 2013. http://www.nsf.gov/about/budget/fy2009/pdf/23_fy2009.pdf.2013. http://www.nsf.gov/about/budget/fy2009/pdf/23_fy2009.pdf.", "previouslyFormattedCitation" : "FY 2009 Budget Request to Congress: Office of Cyberinfrastructure, 2009, http://www.nsf.gov/about/budget/fy2009/pdf/23_fy2009.pdf." }, "properties" : { "noteIndex" : 0 }, "schema" : "https://github.com/citation-style-language/schema/raw/master/csl-citation.json" }
4. Nick W. Jankowski, "The Contours and Challenges of e-Research," in *E-Research: Transformation in Scholarly Practice*, ed. Nick W. Jankowski. (New York: Routledge, 2009), 7.
5. Tobias Blanke and Stuart Dunn, "The Arts and Humanities e-Science Initiative in the UK," in *E-SCIENCE '06: Proceedings of the Second IEEE International Conference on e-Science and Grid Computing* (Washington, DC: IEEE Computer Society, 2006), accessed March 24, 2013. doi:10.1109/E-SCIENCE.2006.261069.
6. Matthew G. Kirschenbaum, "What Is Digital Humanities and What's It Doing in English Departments?," *ADE Bulletin* 150 (2010): 1-7, doi:10.1632/ade.150.XXX.
7. Peter Halfpenny et al., "Developing the UK-based e-Social Science Research Program," in *e-Research: Transformation in Scholarly Practice*, ed. Nicholas W. Jankowski (London, UK: Routledge, 2009), 73-88.
8. *Cyberinfrastructure Framework for 21st Century Science and Engineering (CIF21)*, *NSF-Wide Investiments*, (2012), accessed Faburary 22, 2013. http://www.nsf.gov/about/budget/fy2012/pdf/40_fy2012.pdf.
9. Ibid., 29-34.
10. Research Councils UK & The Royal Society. (2009). *RCUK Review of e-Science 2009: Building a UK Foundation for the Transformative Enhancement of Research and Innovation*, accessed March 22, 2013. http://www.stfc.ac.uk/Resources/pdf/RCUKe-ScienceReviewReport.pdf

11. "National Science Foundation Advisory Committee for Cyberinfrastructure Task Force on Campus Bridging Final Report ," (March 2011), accessed February 20, 2013. http://www.nsf.gov/od/oci/taskforces/TaskForceReport_CampusBridging.pdf.
12. Ellen Collins, "The National Data Centres," in *Managing Research Data*, ed. Graham Pryor (London, UK: Facet Publishing, 2012), 151-172.
13. Sarah Jones, "Research Data Policies: Principles, Requirements and Trends," in *Managing Research Data*, ed. Graham Pryor (London, UK: Facet Publishing, 2012), 47-66."publisher-place" : "London, UK", "title" : "Research Data Policies: Principles, Requirements and Trends", "type" : "chapter" }, "uris" : ["http://www.mendeley.com/documents/?uuid=7023f343-767b-4065-802a-f3dcfc23632b"] }], "mendeley" : { "manualFormatting" : "Sarah Jones, \u201cResearch Data Policies: Principles, Requirements and Trends,\u201d in Managing Research Data, ed. Graham Pryor (London, UK: Facet Publishing, 2012
14. JISC Company, "Managing Research Data Programme 2011-13," accessed February 11, 2013. http://www.jisc.ac.uk/whatwedo/programmes/di_researchmanagement/managingresearchdata.aspx
15. Jisc Company, "Funding," accessed February 11, 2013. http://www.jisc.ac.uk/fundingopportunities.aspx.
16. Mary Auckland, *Reskilling for Research: An Investigation into the Role and Skills of Subject and Liaison Librarians Required to Effectively Support the Evolving Information Needs of Researchers*, (Research Libraries UK, January 2012), accessed February 24, 2013. http://www.rluk.ac.uk/files/RLUK%20Re-skilling.pdf
17. Wendy Lougee et al., *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, andthe Research, Teaching, and Learning Steering*, accessed March 3, 2013. http://www.arl.org/component/content/article/6/1134.
18. Association of Research Libraries,"E-Research," accessed February 14, 2013, http://www.arl.org/rtl/eresearch/escien/index.shtml.
19. Catherine Soehner, Catherine Steeves, and Jennifer Ward, "E-Science and Data Support Services: A Study of ARL Member Institutions" (August 2010), accessed March 3, 2013. http://www.arl.org/storage/documents/publications/escience-report-2010.pdf
20. Carol Tenopir, Ben Birch, and Suzie Allard, *Academic Libraries and Research Data Services: Current Practices and Plans for the Future*, (2012), accessed March 25, 2013. http://www.ala.org/acrl/sites/ala.org.acrl/files/content/publications/whitepapers/Tenopir_Birch_Allard.pdf.
21. Elisabeth Jones, "Reinventing Science Librarianship: Themes from the ARL-CNI Forum," *Research Library Issues: A Bimonthly Report from ARL, CNI, and SPARC*, (February 2009), 12-17.
22. Alex Voss et al., "e-Research infrastructure development and community engagement", in *Proceedings of the UK e-Science All Hands Meeting 2007*. (Nottingham UK: 10-13 September, 2007), accessed March 16, 2013. http://www.allhands.org.uk/2007/proceedings/papers/866.pdf
23. D. Scott Brandt, "Librarians as Partners in e-Research: Purdue University Libraries Promote Collaboration". *College & Research Libraries News*, (June 2007), 365–368.
24. Kristin Briney, "The Lab Notebook: A Critical Tool for Scientific Data Management," slides presented to Data Management Bootcamp, UW-Madison Libraries. (22 February 2012), accessed March 28, 2013. http://www.slideshare.net/kbriney/lab-notebooks-a-librarians-primer
25. Stefan Kramer, "Data Librarianship: Past, Present, Future, Challenges, Opportunities," Invited Presentation to Staff of GESIS (Leibniz Institute for the Social Sciences) in Bonn, Germany (December 2010), accessed March 22, 2013. http://ecommons.library.cornell.edu/handle/1813/19484
26. Alan Masson, "VRE Library Services: Learning from Supporting VLE Users," *Library Hi Tech* 27, 2 (2009): 217-227, doi:10.1108/07378830910968173.