**Nicholas M Weber**

**University of Illinois GSLIS**

**nmweber@illinois.edu**

# Knowledge Infrastructure: The Research Library's Role in Information Transfer

Information products in the networked environment of digital science, or eScience is readily mobile and capable of transfer across vast physical distances. In shifting to a digital medium, eScience now utilizes a data-centric set of methods[1] requiring an infrastructure capable of sustaining long-distance collaborations and high volumes of data (Atkins et al. 2003). This shift towards a digital workspace has also increased the need for both institutions and information professionals to be more involved in the organization, management, collection, and the preservation of this data. In 2003 the National Science Foundation (NSF) recognized that a new form of infrastructure was needed to support the burgeoning practice of digital scholarship and computationally dependent science:

> ...cyberinfrastructure refers to infrastructure based upon distributed computer, information, and communication technology. If *infrastructure* is required for an *industrial* economy, then we could say that *cyberinfrastructure* is required for a *knowledge* economy. (Atkins et al. p. 5)

While a shift in scientific methodologies is readily apparent, this movement from an industrial infrastructure to a cyber infrastructure is not a clear or well-defined progression. Between these two periods there have been a number of changes in how scientists access material, transfer information, disseminate claims and produce knowledge. This essay will attempt to recognize how infrastructural shifts have affected the research library's role in the process of moving data from actor to actor (information transfer) and adding value by means of preservation and intellectual organization (knowledge production). I will survey a small swath of policies and presidential reports concerning scientific information growth in order to contextualize the current milieu of cyberinfrastructure. I then conclude by suggesting that the sociotechnical elements of a research library that supports eScience is ill-suited for cyberinfrastructure discourse, and suggest the application of what I believe is a superior concept: knowledge infrastructure.

## Infrastructure

Infrastructure is most easily recognized as a physical manifestation or a technical standard; railroads, telephone wires, paved roads, textual protocols, and networked technologies are some of the most immediately identifiable examples. However, discussions of infrastructure are troublesome in that a unifying characteristic for all of

---

[1] The methodological shift from a theory and simulation driven mode of inquiry to a computational 'data-centric' practice is often referred to as the fourth paradigm of scientific research (Grey et al 2009).

these examples is hard to accurately define. Paul Edwards has noted that infrastructure is particularly laborious to discuss because of its dependence on technical standards that contribute to much larger, patch-work like structures (2010, p.12). This "patch-work" system seems to be built in layers and has numerous levels of abstraction to further complicate discussions and definitions of infrastructure. Leigh Star's work in this area has recognized a complex component make-up of infrastructure that is "…both relational and ecological -it means different things to different groups and it is part of the balance of action, tools, and the built environment, inseparable from them." (Starr 1999, p. 377) Addressing the idea of component parts, Star and Ruhleder recognized an 'ecology of infrastructure' that has classifiable features (fig 1). Star and Ruhleder's framework of features is particularly useful for discussing how information transfer, although currently magnified in scale, is still essentially facilitated by a core set of infrastructural elements that can be broadly applied throughout the sciences, and wider society. These core features are also useful for identifying the ways in which cyberinfrastructure and knowledge infrastructure either refine or neglect certain elements for the sake of a specific application. Later work by Star and Griesemer also discussed the role of various actors in a research network showing that information management, although centralized, depends on a series of processing[2] (p. 414). As an activity, information processing is readily recognizable in the elements described above, and easily fits within

| **Feature** | **Distinction** |
|---|---|
| Embededness | Operates within social structures and technological deployments |
| Transparency | Often invisibly supports tasks |
| Beyond single event or local instance | Served by a continuum, transcending physical boundaries |
| Learned as part of membership | It's members or artifacts take it for granted is an attribute (new members go through period of acclimation) |
| Links with Conventions of Practice | Shapes and is shaped by the conventions of practice. |
| Embodiment of Standards | Infrastructures takes on transparency by plugging into other infrastructures and tools in a standardized fashion |
| Built on Installed Base | Inherits strengths and limitations of constituent systems |
| Becomes Visible upon Breakdown | As a result of its far reaching effect |
| Fixed in modular increments | Infrastructure is not completed all at once or globally (never changed from above because of the component parts) Changes require time, negotiation, |

Figure 1. Star and Ruhleder's features and distinctions of infrastructure.

---

[2] This chain of information work undoubtedly transcends the digital landscape; Peter Burke notes that in Europe as early as the seventeenth century, information traveled from the periphery to economic centers by means of an "assembly line" of information processing (p. 75).

the common conception of infrastructure as the "plumbing" or "pipes" for information transfer. What is much less clear is how the actors in a research network that facilitate knowledge production are accounted for in their manipulation of technology to, "…share and maintain specific knowledge about the human and natural worlds" (Edwards 2010, p. 12). The research library has traditionally played this role in academic settings, but eScience and its increasingly large production of data has caused a major recalibration of this activity.

## Information Growth and Infrastructural Planning

The research library and indeed the field of library and information science have been struggling with unwieldy amounts of scientific information from their very inception. Courses in indexing, abstracting and classification of scientific material were amongst the first to be offered in programs focusing on library science (Buckland 1996). In England, as early as 1894 H.E Armstrong (echoed by JG Pearce in 1924) was calling scientific periodicals "unmanageable" in terms of both size and growth (Muddiman, p.57). The same concept was being addressed in the United States at conferences like "The Problem of Specialized Communication" in 1952 and a symposium to address the growth biological communications in 1960 (Conrad et al. 1961). The latter symposium painted a picture of biological publications in a state of crisis so dire that Foster Mohrhardt likened the situation to a tower of babble (p. IX) At the University of Chicago's specialized communications conference, Verner Clapp warned of the big business that was scientific publication, but ultimately had a deft condemnation for those proclaiming the rise in publication to be a revolution:

> The problems of the communication of specialized information in modern society are little more than a numerical multiplication of the factors of identical problems as they have existed at any time since the communication of specialized information became a recognizable function of society (p. 12-13)

This sensibility was echoed by many in the library and scientific publishing world during the 1950's and 60's, just as numerous plans to combat the scale of information transfer were envisioned by the likes of Bush's Memex machine and Licklider's "connected desk" (Segaller, 1998). These also included educational initiatives for future practitioners, such as Mortimer Taube's offering of the first course in documentation at Columbia in 1951 (Taylor, 1976). Likewise, the bibliometric world was just beginning seminal work to measure information transfer by calculating scientific citation practices, evaluating impact and estimating half lives (Price, 1963; Garfield and Sher, 1963). Beyond the loose coupling of these activities to LIS, bibliometric studies formalized many of the assumptions made about what and how scholarly scientific literature was being produced and consumed. Specific to the coordination of library service, L.J.B Mote conducted one of the first comprehensive studies identifying user needs and service models in a framework consumable by numerous types of information professionals (1976).

Events and publications such as the ones mentioned above display a field grappling with increased information volumes in sophisticated ways. Ultimately, the overwhelming sense of increasing volumes of information were tempered by case studies, user analysis and theoretical frameworks that reconceived how research could best be supported. Thus, the infrastructure of knowledge production has been recalibrated numerous times in the paradigm shifts leading up to the current data-centric practice of eScience.

As both Star and Edwards alluded, infrastructure is difficult if not impossible to talk about without some discussion of the ancillary parts that were assembled to make it so. In order to more fully understand the contemporary role of the research library it is useful to consider historical visions of what information transfer required, and how these align with current cyberinfrastructure initiatives. A majority of past planning efforts for infrastructural development is found within technical reports and committee briefings at a national level. This is due in no small part to the almost exclusive role that the national government of the United States plays in funding science research and passing legislation to affect large infrastructural change (Stokes, 1997).

## Infrastructure Policy and Planning: The Research Library

Libraries and their technological components were piecemeal in early infrastructural support of scientific information. One of the earliest examples might be the 1934 Bibliofilm venture by the USDA library that attempted to lend microfilming technologies to research publications (Schultz and Garwig, 1969). There were also considerable discussions about the evolution of publication models and delivery methods during this period, but must notably were the grand proposals for a national science foundation to presidential advisory committees. Two of these seminal proposals, “Science- The Endless Frontier” by Vannevar Bush in 1945, and it’s counterpart “Science and Public Policy” (often referred to as the Steelman Proposal) in 1947, suggested centralized funding and organization of science by means of a national program (Blanpied, 1999). Perhaps most notably, both proposals attempted to define the infrastructural elements necessary to support, renew and harness the flow of scientific information emanating from US based research, which was steadily increasing in both complexity and volume post World War II (Pinelli et al. 1992). Although neither proposal was immediately successful, they had a profound impact on future infrastructural proposals in terms of both the vision and the scope of an appeal one could make to the president’s committee (Pinelli et al. 1992). Both proposals however lack any substantial discussion about the preservation, organization or dissemination of the products of basic science research. In fact, the word library is not found in either document.

The subsequent foundation of a National Science Foundation (NSF) in 1950 did little to rectify the shortcomings of earlier proposals in terms of an infrastructure for the organization and preservation of research products. It wasn’t until 1958 in a report on the expanding informational needs for scientific research that William Baker addressed the importance of organizing and providing meaningful access to the products stemming from NSF funding (1958). In this report Baker notes the Library of Congress increased holdings becoming untenable and the mounting reports available from completed NSF funded projects unsustainable (p. 3). His proposal was a scientific information service that could be offered either centrally, or in a distributed system (p. 9). His supporters sided with the former model, and the Office of Science Information Service was established by NSF in later that year (Pinelli et al. 1992).

The landscape of technological innovation was rapidly expanding, but infrastructural directions for the country were much less clear for policy makers in the 1950’s and 60’s.

Visionaries like Vannevar Bush were who had once proposed (at the time) seemingly crackpot desktop computational abilities were being revisited in the wake of computing breakthroughs like stored program retrieval and solid state circuitry present in second generation computers (Emard, 1976). Additionally, these hardware breakthroughs were contributing a fuller vision of what networked scientific work might realize and what the necessary infrastructural components would require (beyond the programmable computer). Like Bush's Memex proposal, J.C.R Licklider's "Library of the Future" included a desk that might connect to a central system for sharing and distributing knowledge. Licklider however recognized that the electronic "umbilical cord" (Licklider 1965, 33) necessary to connect his proposed hardware was likely the most important collaborative development necessary to advance information transfer[3]. It's worth noting that Licklider's "vision" was imagined under the auspice of a science library and that the infrastructural elements were both organized and expertly serviced in this environment.

This decade also saw a powerful and sweeping report filed to the president's science advisory committee by Alvin Weinberg entitled "Science, Government, and Information" (1963). Weinberg believed that a recalibration of research attitudes and practices was needed, and that this social component should be initiated by a central authority. He wrote famously that "…the attitudes and practices toward information of all those connected with research and development must become indistinguishable from their attitudes and practices toward research and development itself" (p. 17). Weinberg outlined a "crisis" in scientific communication, acknowledging that the infrastructural elements necessary to interconnect many of the disparate parts were effectively beyond central control, "because these communication systems have grown up in isolation, they too often tend to further fragment our already disjointed scientific structure" (p. 10).

Particularly pertinent to a discussion of infrastructure Weinberg recognized that communication systems are simply a component part of a larger mechanism that should include the management, storage and dissemination of scientific research (p. 13). His ultimate recommendation though distanced libraries and librarians (in name) from this process, "…the specialized information center should be primarily a technical institute rather than a technical library. It must be led by professional working scientists and engineers who maintain the closest contact with their technical professions…" (p. 6) Expanding on this thought he later more directly states that, "Communication cannot be viewed merely as librarians' work; that is, as not really part of science. An appreciable and increasing fraction of science's resources, including deeply motivated technical men as well as money, will inevitably have to go into handling the information that science creates" (p. 17). However, throughout this report there is mention of the need for a "middle-man" for science information processing, and particularly a "documentalist" with domain knowledge. The need for knowledge workers or information professionals with domain knowledge is a prime example of the complexity of infrastructure and the unintended consequences of its reshaping. The retooling of an information workforce also requires a shift in traditional avenues of employment, educational models, cultural acceptance and even a theoretical base, the likes of which LIS programs are still

---

[3] This proposal would be later realized by ARPANET and DARPA /NASA as the internet ( Segaller 1998).

grappling with today (i.e. Gold 2008; Cragin et al. 2007).

In 1976 NSF commissioned report titled “A national approach to scientific and technical information in the United States.” The author, Joseph Becker stated that, “…the hardware and software tools that accompany most of our operating science information systems are applied and managed, by and large, by information scientists”(p.12). Becker also recognized the actors of a scientific information transfer infrastructure were varied, but included both corporate and academic / research libraries (p.27). These small acknowledgements were amongst the first nationally to make explicit the burgeoning role of information scientists and research libraries in the management of data rather than the passive collection and organization of periodicals. In many ways Becker’s report foregrounded the issues leading up to NSF’s conceptualization of cyberinfrastructure, most notably that science was now (in 1976) producing “…staggering quantities of data for analysis. Interpretation, and retrieval…New approaches to science information wholly different from the classical systems used to process publications will be required to handle the data efficiently.” (p.20). Becker also lauded the potential of a networked infrastructure that included remote sites in addition to libraries and information centers (p. 37)

These networked connections were being realized at a national level by the early 1990’s and infrastructure funding would soon follow from NSF, DARPA and NASA. The Digital Libraries Initiative announced in 1994 granted six separate awards to test and demonstrate new technologies in developing digital communication networks (Griffin, 1998). Perhaps as important as any research stemming from this funding were the digital library workshops[4] in which infrastructural definitions of a digital library were redefined and renegotiated beyond an initial institutional affiliation (Griffin 1998).

In the same year, the Association of Research Libraries and the American Association of Universities convened a taskforce to investigate a ‘National Strategy for Managing Scientific and Technical Information’ This report viewed academic libraries as a component part in the larger infrastructure being created by the US Government, saying “…It is important for universities to participate in the development of such a system so that it reflects both the research and educational needs of those institutions.” (AAU, 1994) The recommendations from this report included a new role for libraries in the management of scientific data and a lengthy discussion of the future for disseminating those holdings to a network of institutions. Though these findings are hardly different from the reports offered by Becker some fifteen years earlier, the ARL report increased the academic libraries focus on incorporating existing schemas for information transfer, and showed a heightened awareness of an emerging national infrastructure that could be harnessed for higher education’s benefit. This report also signifies an important shift in the research library’s role in building infrastructure and clearly articulates the need for

---

[4] The 1995 IITA Digital Libraries Workshop entitled "Interoperability, Scaling, and the Digital Library Research Agenda,” http://www.ccic.gov/pubs/iita-dlw, and the NSF sponsored Santa Fe planning workshop on distributed knowledge work environments [http://www.si.umich.edu/SantaFe/]

practitioners to be involved in the development of technical standards and protocols that will enable future scientific research.

In 1999, an NSF report to the president began to articulate future needs for both the education and technical infrastructure to sustain a flourishing digital economy.
The report concludes that research funding, especially that of information infrastructure was wholly inadequate to sustain the types of advances achieved in computing over the last decade (NSF 1999, p.10). Throughout, this report stresses the importance of research in information technology to spur further economic expansion, and specifically calls for a 500 percent increased in ‘scalable information infrastructure’ over a five year period. In terms of sustainability, the report makes strong recommendations about increasing the duration of project funding period and an increase in funding to research long-term preservation strategies(NSF 1999, p. 12) One of the key recommendations for infrastructure is also the development of middle-ware which enables large scale system integration (NSF 1999, p. 12). This particular recommendation articulates an early conceptualization of cyberinfrastructure for eScience research and hints towards future funding initiatives for collaborative work.

## Cyberinfrastructure and the Library

The intersection of computing, information needs and communication technologies is an area well understood by academic libraries and indeed central to the field of LIS research. However, early in the twenty-first century research libraries were being overwhelmed by the proliferation of computing power in eScience. Until recently terabytes were a scale not well accommodated by libraries as noted by an ARL report in 2007, “…although technology capacity in libraries has grown considerably in recent decades, it is not of the scale or complexity of the e-science environment.” Furthermore, data science and data management is hardly understood by practicing librarians more accustomed to bibliographic material and electronic aggregations of serialized journals (Gold 2008). Libraries themselves have only recently started to collect and preserve the electronic texts of their scholars in the form of institutional repositories (Lynch 2003).

| | Traditional Library | eScience library |
|---|---|---|
| Structural Organization | Discipline based, | Topical research; Cross domain; organized around methodologies |
| Digital Management | Text and stable images | Tabular Data, arrays, spectra, Video or Animated Sequences. |
| Users | Direct affiliation; geographic proximity. | Indirect, remote collaborations that cross-institutional borders |
| Licensing | Negotiation with vendors; tied to object use rather than object collection | Increasingly cross-boundary, and personal claims to information object ownership. |

Figure 2. Key differences in library support for eScience

The success of cyberinfrastructure will depend in part on a retooling of the existing LIS workforce. This sort of recalibration in support of new scientific methods has historical roots not just in Becker's 1976 report to NSF, but all the way back to 19th century. David Muddiman notes that infrastructural changes aimed at "replacing voluntarism and individualism" from 1870 onwards included new science education models, such as the establishment of natural science faculties and many traditional discipline based departments still present in many universities today (2009, p. 55). This same recalibration can be seen taking place since the early 2000's as informatics programs have steadily spread throughout the departmental make-up of higher education. LIS is no exception, many of the top ranked programs offer a disciplinary focus on data or digital curation which might be considered the service component of eScience research (Cragin et al. 2007) or cyberinfrastructure (NSF 2007).

However, these sorts of adjustments and the new curatorial activities performed by this retooled workforce are not cyberinfrastructure activities. This is a crucial and often overlooked distinction: Cyberinfrastructure as it's been defined and described by NSF (Atkins et al. 2003, NSF 2007) acknowledges the need for new education and service models, but it does not include any of the social element necessary to both capture and sustain knowledge production. Its clear that by enabling new means of information transfer, data sharing, visualization and reuse science will be more capable of meeting contemporary research challenges. But these opportunities are not socially accounted for in cyberinfrastructure.

What's needed is another level of infrastructure abstraction to accurately identify the elements of knowledge production and preservation that a research library, as an institution can sustainably provide. This abstraction needs to accommodate for both the curation activities that support a lifecycle model of data as well as the burgeoning role of information professionals beyond a simple "liaison" status. Others have suggested that sociological, and sociotechnical studies are worthy investigations for cyberinfrastructure research (Lee et al. 2006, Ribes and Lee 2010). I believe that a coupling of the social elements in infrastructure with cyberinfrastructure is mistaken and overestimates the capabilities of a cyberinfrastructure framework.

## Knowledge Infrastructure

Knowledge infrastructure, as evinced by Paul Edwards includes "…networks of people artifacts and institutions that generate, share and maintain specific knowledge about the human and natural worlds." (p.12). Knowledge infrastructural elements include entities at an individual level and institutional level that are networked for both practical and theoretical collaboration. Knowledge infrastructure also allows for a more accurate framing of curatorial activities that enhance existing information transfer activities, and recognizes that knowledge can be both shared and maintained through these technological networks. Edwards also notes that, "knowledge infrastructure is not a new concept in science, it is often discussed in terms of 'tehchnoscience' to capture the technological dimension of science as knowledge practice." (p.19) This is easily identified in the literature surrounding social dimensions of cyberinfrastructure that mistake the middleware systems, for the "middleman" service components. Knowledge infrastructure also accounts for the preservation and persistence of information that is necessary for eScience to sustainably grow. Edwards call knowledge infrastructure "a superior concept" to other infrastructure frameworks that try to incorporate a disjoint sociotechnical dimension, "…because it considers endurance, reliability, and the taken for grantedness of a technical and institutional base supporting everyday work and action." (p.19) The institutional base that provides these elements is unquestionably a research library.

With knowledge infrastructure as a conceptual framework researchers in LIS are able to answer questions that move from investigating ability to measuring effectiveness. Research questions such as: Do data repositories facilitate better data discovery for interdisciplinary researchers than traditional informal sharing? Are datasets and resulting publications linked in ways that facilitate reuse and reanalysis? How are data licensing policies by institutions inhibiting large-scale meta-analyses? Are open-access policies issued by funding mandates effectively and sustainably enforced? These are questions natural scientists are both unequipped, and incapable of effectively answering about their communities of practices. But these questions are crucial in measuring the impact of cyberinfrastructure funding, and necessary to expose gaps in infrastructural capability for both eScience researchers and funders.

## Information transfer to knowledge production.

I have attempted to show the various components of cyberinfrastructure and knowledge infrastructure as a means of better understanding the general infrastructural framework for research libraries facilitating eScience activities. The academic research library has historically played a preservation and organization function in the process of information transfer and knowledge production. These functions are evolving and changing dramatically in the face of eScience research. As demonstrated by the review of historical policy and technical reports of the United States, infrastructural planning for science often identifies the needs of information transfer through the growth of scientific output, but rarely consider the sustainability of the solutions proposed. Cyberinfrastructure is no exception. Its funding mandates allow for new complex systems to be built, interoperability of platforms to be negotiated and data products to be meaningfully and accurately exchanged. However, to make use of data or to produce knowledge data needs to be normalized, described and organized in meaningful ways. Cyberinfrastructure researchers such as Hey and Trefethen have even argued that access, integration and curation of data are as important as storage and computing facilities (2005). Knowledge infrastructure provides these important components, but does so by means of adding an additional layer of services that are discernable and separate from cyberinfrastructure. If cyberinfrastructure is "required for a knowledge economy," then we might say that knowledge infrastructure is a required for sustainable economic growth.

## Works Cited

Association of Research Libraries. (1994). Reports of the AAU Task Forces on Acquisition and Distribution of Foreign Language and Area Studies Materials, a National Strategy for Managing Scientific and Technological Information, Intellectual Property Rights in an Electronic Environment. Washington, DC: Association of Research Libraries.

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, Nov. 2007. http://www.arl.org/bm~doc/ARL_EScience_final.pd

Atkins, D. et al., Eds. (2007) Report of the National Science Foundation Blue- Ribbon Panel on Cyberinfrastructure, Revolutionizing Science and Engineering Through Cyberinfrastructure, (http://www.nsf.gov/od/oci/reports/atkins.pdf).

Baker, W. O. (1958). Improving the Ability of Scientific and Technical Information in the United States.

Becker, J. (1976). A National Approach to Scientific and Technical Informaiton in the United States. Washington D.C.

Beagrie, Neil (2006). Digital Curation for Science, Digital Libraries, and Individuals. International Journal of Digital Curation, 1(1), 3-16. http://www.ijdc.net/ijdc/article/view/6/5.

Blanpied, William A. (1999). Science and Public Policy: The Steelman Report and the politics of post-world war II science policy. Science and Technology Yearbook, AAAS c. 29 http://www.aaas.org/spp/yearbook/chap29.htm

Buckland, Michael (1996). Documentation, information science, and library science in the USA, *Information Processing and Management*, Vol. 32, No. 1, pp. 63-76.

Bush, V. (1945) Science - The Endless Frontier. Washington, D.C.: United States Government Printing Office.

Burke, P. (2000). Locating Knowledge: Centres and Peripheries. A Social History of Knowledge: From Gutenberg to Diderot. Polity.

Clapp, V. W. (1952). THE COMMUNICATION OF SPECIALIZED INFORMATION Papers Presented before the Seventeenth Annual Conference OJ the Graduate Library School OJ the BY.

Conrad, G. Miles , Conway Zirkle and Paul B. Sears (1961) *Public Health Reports (1896-1970)* Vol. 76, No. 11, pp. 1034-1036

Cragin, Melissa H., P. Bryan Heidorn, Carole L. Palmer, and Linda C. Smith. An Educational Program onData Curation. Poster presented at the Science and Technology Section of the annual American Library Association conference, Washington, D.C., June 25 2007.

Edwards, P. N. (2010). A vast machine: Computer models, climate data, and the politics of global warming. Cambridge, Mass: MIT Press.

Emard, J. P. (1976). An information science chronology in perspective. Bulletin of the American Society for Information Science , 2(8), 51-56

NSF (2007). Cyberinfrastructure Vision for 21st Century Discovery. (March 2007).

Garvey, W. D. (1972). COMMUNICATION AND INFORMATION PROCESSING WITHIN SCIENTIFIC DISCIPLINES : EMPIRICAL FINDINGS FOR PSYCHOLOGY, 8, 123-136.

Gabridge, Tracy (2009). The Last Mile: Liaison Roles in Curating Science and Engineering Research Data. Research Library Issues: A Bimonthly Report from ARL, CNI, and SPARC, no. 265 (August 2009): 15–21. http://www.arl.org/resources/pubs/rli/archive/rli265.shtml.

Garfield, E., Sher, H., 1963. New factors in the evaluation of scientific literature throughcitation indexing. American Documentation 14, 195–201.

Gold, A. (2007). Cyberinfrastructure, Data, and Libraries, Part 2. D-Lib Magazine 13(9/10). Retrieved November 17, 2008, from http://www.dlib.org/dlib/september07/gold/09gold-pt2.html

Gray, J., Szalay, A. S., Thakar, A. R., & Stoughton, C. (2002). Online Scientific Data Curation , Publication , and Archiving One Microsoft Way Online Scientific Data Curation , Publication , and Archiving. Building, (July).

Griffin, S. M. (1998). NSF / DARPA / NASA Digital Libraries Initiative A Program Manager ' s Perspective. D-Lib Magazine, (September).

Heidorn PB: Shedding Light on the Dark Data in the Long Tail of Science. Library Trends 2008 , 57:280-289.

Hey, T., & Trefethen, A. E. (2005). Cyberinfrastructure for e-Science. Science (New York, N.Y.), 308(5723), 817-21. doi: 10.1126/science.1110410.

Lee, Charlotte P., Dourish, Paul and Mark, Gloria (2006) 'The human infrastructure of cyberinfrastructure', Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work NY, NY: 483-92.

Mohrhardt, S. (1961). Symposium on Biological Communications was part of the dedication program for Biological Abstracts' new headq uarters building on October. Biological Abstracts.

Muddiman, D. (n.d.). Part ii infrastructure, networks and the state. Networks.

NSF. (1999). Information Technology Research: Investing in our Future. Report to the president. Retrieved from http://itrd.gov/pubs/pitac/index.html.

NSF. (2007). A Cyberinfrastructure Vision for 21st Century Discovery. (March 2007).

Licklider, J.C.R. (1965). Libraries of the Future. Cambridge: The MIT Press.

Lynch, C.A. (2003). Institutional Repositories: Essential Infrastructure for Scholarship in the digital age. ARL: A Bimonthly Report on Research Library Issues and Actions from ARL, CNI, and SPARC, Issue 226 (February 2003)

Mote, L. J. B. (1962). Reasons for the Varitions in The Information Needs of Scientists. Journal of Documentation, 18(4).

Pinelli, T. E., Henderson, M., Bishop, A.P., & Doty, P. Chronology of Selected Literature, Reports, Policy Instruments, and Significant Events Affecting Federal Scientific and Technical Information (STI) in the United States, 1945-1990, (Washington, DC: National Aeronautics and Space Administration, NASA/DoD Aerospace Knowledge Diffusion Research Project, Report Report No. 11 (NASA TM-101662)).

Price, D. J. S. (1986). Little science, big science-- and beyond. New York: Columbia University Press

Ribes, D., & Lee, C. P. (2010). Sociotechnical Studies of Cyberinfrastructure and e-Research: Current Themes and Future Trajectories. Computer Supported Cooperative Work (CSCW), 19(3-4), 231-244. doi: 10.1007/s10606-010-9120-0.

Schultz, C. K., & Garwig, P. L. (1969). History of the American Documentation Institute--A sketch. American Documentation , 20(2), 152-160.

Segaller, S. (1998). Nerds: A Brief History of the Internet. TV Books, New York. (quote by L. Roberts.)

Soehner, C., Steeves, C., Ward, J,. (2010) E-Science and Data Support Services: A Study of ARL Member Institutions. Association of Research Lbiraries. http://www.arl.org/bm~doc/escience_report2010.pdf

Star, S. L. (1999). The Ethnography of Infrastructure. American Behavioral Scientist, 43(3), 377-391. doi: 10.1177/00027649921955326.

Star, S. L., & Ruhleder, K. (1996). Steps Toward an Ecology of Infrastructure: Design and Access for Large Information Spaces. Information Systems Research, 7(1), 111-134. doi: 10.1287/isre.7.1.111.

Star, Susan Leigh, Griesemer, J. R., Studies, S., & Aug, N. (1989). Institutional Ecology, Translations and Boundary Objects : Amateurs and Professionals in Berkeley 's Museum of Vertebrate Zoology , Museum, 19(3), 387-420.

Stokes, Donald E., 1997: Pasteur's Quadrant: Basic science and technological innovation, Brookings Institution Press.

Taylor, R. S. (1976). On education. Bulletin of ASIS , 2(8), 34.

Weinberg, W.O. (1963). Science, Government, and Information: The Responsibilities of the Technical Community and the Government in the Transfer of Information. Wasington D.C.