Health Sciences Documentation and Networked Hypermedia: An Integrative Approach

by ANNE J. GILLILAND-SWETLAND

Résumé

L'auteur étudie les possibilités et les limites du réseau Internet pour la diffusion de documents historiques relatifs aux sciences de la santé. Elle fonde ses observations sur les expériences du *Michigan Digital Historical Initiative (MDHI)*, un projet de numéralisation à l'échelle de l'État en histoire des sciences de la santé. Elle définit ensuite le concept de «bibliothèque numérisée» et évalue ses possibilités d'intégration et d'accessibilité croissante, tant physique qu'intellectuelle, à la documentation concernant les sciences de la santé. Elle soutient que le concept de cohérence numérique, tel qu'employé dans les environnements de réseaux hypermedia, tels que le World Wide Web et les bibliothèques numérisées, peut conduire soit à une augmentation ou à une fragmentation du contexte pour le contenu numérique. Que le contexte soit augmenté ou fragmenté dépend du niveau de planification intentionnel au niveau de la conception du système et du degré d'inclusion des documents archivistiques et des descripteurs. L'auteur développe ensuite une vision de la fonctionnalité et de la conception nécessaires à l'inclusion efficace du contenu historique dans les bibliothèques numérisées.

Abstract

The author examines the capabilities and limitations of World Wide Web home pages for disseminating historical materials in the health sciences. She bases her observations on the experiences of the Michigan Digital Historical Initiative (MDHI), a statewide collaborative digital project in the history of the health sciences. She then defines the concept of the "digital library" and examines its potential for integrating and increasing physical and intellectual access to health sciences documentation. She asserts that the concept of digital coherence as applied in networked hypermedia environments such as the World Wide Web and digital libraries can lead either to augmentation or to fragmentation of context for

digital content. Whether context is augmented or fragmented depends upon the level of conscious planning that goes into system design, and the extent to which actual archival materials as well as descriptive surrogates are included. The author then develops a vision of the functionality and design necessary for the effective inclusion of historical content in digital libraries.

Introduction

> The man who classifies facts of any kind whatever, who sees their mutual relation and describes their sequence, is applying the scientific method and is a man of science. The facts may belong to the past history of mankind, to the social statistics of our great cities, to the atmosphere of the most distant stars, to the digestive organs of a worm, or to the life of a scarcely visible bacillus. It is not the facts themselves which form science, but the method in which they are dealt with. The material of science is co-extensive with the whole physical universe, not only the universe as it now exists, but with its past history and the past history of all life therein. When every fact, every present or past phenomenon of that universe, every phase of present or past life therein, has been examined, classified, and co-ordinated with the rest, then the mission of science will be completed.[1]

In 1911, biometrician and geneticist Karl Pearson wrote of how examination, classification, and coordination of facts support the mission of science and in a broader sense bring about new insights and the development of new knowledge. Networked hypermedia exploit computer capabilities to relate and disseminate materials in multiple media in ways not previously possible, and add a new dimension to the conduct of science and the processes of knowledge development in general. The underlying premise of this assertion is that individuals, collaborating groups, and institutions can use network communications technologies coupled with digital hypermedia to create, examine, classify, coordinate, manipulate, and disseminate information and documentary resources of all types without spatial and media constraints.

This paper outlines a vision for an integrated research and administrative environment constructed with networked hypermedia for contemporary and historical health sciences materials of all types. It examines how the most popular manifestation of networked hypermedia, the World Wide Web home page, falls short of realizing that vision. The paper discusses the development and experiences of the Michigan Digital Historical Initiative, an application in the history of the health sciences that is currently available on the World Wide Web, as an example of what a home page can currently achieve, and where it has limitations. The paper then defines the concept of the digital library and discusses its potential for attaining this integrated vision by changing how health sciences documentation is created, associated, contextualized, disseminated, and used. The paper concludes with a discussion of the functionality and design necessary for the effective inclusion of historical content in digital libraries.

The Potential and the Limitations of Existing Hypermedia Environments

The term "networked hypermedia" as used in this paper refers to digital files of interactive multimedia that are linked by hardcoded, non-linear structures, and that can be accessed by and transferred to any site that is networked to the computer site on which the files reside.[2] Multimedia could include text, numeric, image, graphics, video, or audio files either of materials that were originally created digitally, or that have been rendered as digital facsimiles.

Imagine a universe of linked digital resources relating to health care and the health sciences that would allow archivists and manuscript curators, museum professionals, librarians, and medical records administrators to collate virtually collections that are physically divided between repositories; build structural links to map data elements contained in both contemporary and historical health records; and develop digital documentation plans that take into consideration the entire "historical record" (including health care and scientific instrumentation, scientific data, even historical sites). This universe is also able to facilitate electronic ordering, validation, delivery, and payment for digital content; maintain detailed use statistics; and even recoup digitization and system development and maintenance costs through the sale of secondary products such as extracted clinical or scientific data or educational packages of digitized historical materials.

Imagine a universe that would allow:

- Art historians to map images created by computed axial tomography (CAT) and magnetic resonance imaging (MRI) scans, or digitized cryosections of the human body, onto digital facsimiles of the work of daVinci and Vesalius, to establish anatomical accuracy.
- Cardiology students to watch digital video presentations of heart catheterization while listening to an oral history of a cardiologist discussing the development of the procedure.
- Microscope collectors to use trade catalogues and artifactual images to trace the development of the instrument from the work of Galileo through today's probe microscopy.
- High school students to learn the stories of early minority pioneers in healthcare fields; or manipulate digital representations of DNA code, chromosomes, and genes after studying James Watson's discovery of the double helix.
- Public health officials to review demographics on reportable diseases extracted from contemporary and historical reports, and incorporate historical literature and radio broadcasts into current public health campaigns.
- Practitioners and scholars of alternative medicine to locate product sources, and learn of non-allopathic applications and techniques.
- Geneticists to study inherited characteristics by examining four hundred years of birth, death, and marriage records of a closed religious community.

There are some technology pundits who would argue that this universe already exists in the part of the Internet known as the World Wide Web (WWW). The WWW

allows for the rapid construction of, and network access to, hypermedia resources known as "home pages." Using a mark-up language called Hypertext Mark-up Language (HTML) and an Internet computer server, any individual or institution can create and make available a home page containing creator-defined links to digital files that reside on computers at almost any networked site. Indeed, home pages that might be used in each of the use scenarios laid out above are already available.[3] Some home pages, such as those of the National Library of Medicine's History of Medicine Division, or the University of Michigan-based Michigan Digital Historical Initiative, are quite substantial in terms of the amount of historical and supplementary reference materials that are included. The History of Medicine Division page, for example, contains almost sixty thousand images from the Division's prints and photographs collections. Others amount to little more than a hyperlinked guide to further Internet resources.

Home pages have limitations, however, that are difficult to surmount in an environment where creators are not constrained by technical, administrative, descriptive, or design standards or conventions, nor well-informed by empirical knowledge about user groups, needs, or expectations. These limitations could be grouped under the rubrics of *structure and design*, *content*, and *validation*.

Structure and design: Home pages, whether created by archival repositories or clinical facilities, are frequently developed without either a well-developed management strategy or a consciously designed structure. As a result, those who maintain home pages find it difficult to continue to expand them without resorting to labyrinthine manoeuvres and a billboard of navigation icons. Even where an institution does develop its own design protocols and system structure, this breaks down the moment they link to the home page of another institution which does not adhere to the same conventions. Users are confounded by a lack of systematic browsing trajectories and retrieval procedures, and when they do find something they would like to view, it is too often delivered without adequate contextualization or explanation.

Content: Home pages lack depth of content and breadth of coverage because in many cases they contain predominantly only descriptions, rather than digital versions of the actual materials. This is in part because many of those descriptions (e.g., catalogue records or finding aids) already exist as digital text and can quickly be marked up and loaded on the Web. In greater part, it is because so much of the world's documentation still exists only in hardcopy, often also with significant legal, ethical, and economic constraints on how much can be digitized, reproduced, and generally disseminated. As a result, many repositories have only digitized and mounted those materials that are legally and ethically reproducible, those for which they have received special funding, or those that are representative of wider collections—all in all, not a very systematic documentary approach, and certainly one that takes neither the array of related resources potentially available from other institutions nor user needs into account.

Validation: Validation in the digital environment has three major components. The first component is establishing that the source that is delivering the digital content is reliable. Many home pages are created by individuals

> acting as their own publisher or information clearinghouse without an institutional imprimatur. The second component is assuring users that the digital content being delivered is complete and accurate. This is difficult to achieve without a standard structure that incorporates much more metadata about digital content than currently exists. For example, users generally are not informed of how content was selected for digitization and inclusion in the home page, what other material exists but remains undigitized and unavailable online, nor how content might relate to other materials within both the same fonds and other collections. Moreover, without standard digitization procedures that clearly state if and how any content may have been altered during the digitization process, it is not possible to demonstrate content completeness and accuracy to users. The third component comes into play when a user enters into a privileged transaction with a digital environment (for example, ordering and paying for images or accessing confidential data) and must provide proof of his or her identity. This component is an essential one for a health sciences environment where materials may have very specific access constraints. It is also essential, however, if any form of commercial transaction is to take place on the World Wide Web. As a result, this limitation is being addressed by many parties besides archivists.

Taken together, such shortcomings severely limit the potential of these home pages to achieve the visions articulated above, and result in the dissemination of documentary content that has been removed from meaningful context. In other words, the use of hypermedia to build links between associated digital materials without taking structure and design, content, and validation into account, rather than augmenting the context that archivists have carefully preserved for those materials, has resulted in fragmenting that context without accruing any real advantage over traditional methods of conducting archival research and reference.

WWW home pages, therefore, while they may whet the user's appetite for an integrated reference and research environment, mostly highlight the need for more comprehensive, reliable, and systematically presented digital resources. In the words of Clifford Stoll:

> ... for all of this communication, little of the information is genuinely useful. The computer gets my full attention, yet either because of content or format, the network doesn't seem to satisfy.[4]

The Michigan Digital Historical Initiative

The Michigan Digital Historical Initiative (MDHI), as one of the earliest and today one of the most wide-ranging digital applications in the history of the health sciences, provides a good case study of the potential and limitations of networked hypermedia as conceived of in home pages. MDHI grew out of the W.K. Kellogg-funded SourceLINK Project at the University of Michigan Historical Center for the Health Sciences. SourceLINK was funded in 1993 to develop an electronic clearinghouse and consulting service for statewide archival resources in the history of the health sciences in Michigan. The project grew rapidly due to a high level of participation by Michigan archives and museums and the University of Michigan School of

Information and Library Studies, and a growing awareness of the potential of digital multimedia and network technology. As a result of this growth, MDHI, an informal consortium of several Michigan historical repositories, as well as educators, technologists, and health science professionals, was formed, and its World Wide Web home page first went up in March 1994.

MDHI is coordinated by the Historical Center for the Health Sciences but has broader objectives than those conceived for the original SourceLINK Project.[5] Before building the MDHI World Wide Web environment, MDHI participants undertook a series of planning discussions that resulted in the development of a vision of what MDHI should seek to achieve, and a series of assumptions upon which that vision was predicated.

Vision:

- Develop a statewide demonstration project that could serve as a feasibility study and model for similar work with other subject areas.
- Build an integrated digital environment in a manageably-sized and topical subject area. This environment could build virtual links between multiple archival and museum collections, media, and collection surrogates that are provenancially or intellectually related but are physically disassociated from each other due to historical, professional, or disciplinary divides. Such an environment would, in effect, augment the limited contextual information generally available to a user from any one repository source. For example, the personal papers of a prominent surgeon split between the archival collections of two universities with which the surgeon had been associated could be virtually reunited (either by relating two finding aids, or by concatenating digitized versions of the actual collection). Moreover, logical links could be built to additional historical materials that might have a bearing on developing an understanding of the medical contribution of that surgeon, such as images and descriptions of instrumentation he or she developed for innovative surgical procedures, patient records tracking the outcome of such procedures, or published analyses of surgical developments of that time.
- Since traditional cataloguing and preservation of collections do not necessarily make them accessible to all audiences, this project would help to overcome some of the need to travel to actual repositories to view unique historical resources, and also help to make those resources more intellectually accessible.
- Support a proactive emphasis on marketing historical resources to meet contemporary needs by identifying and building user markets for digital resources, while at the same time balancing this against a need to consider a democracy of access across different user groups.
- Explore structural and presentation issues such as indicating and maintaining the relationship of digital materials within their traditional descriptive hierarchies, automatic linking to additional associated digital objects, screen design, multiple navigation mechanisms, and the most effective ways to present and disseminate images.
- Investigate and address security issues such as document and user authentication, privacy of record subjects, and automatic financial transactions.

Assumptions:

- The project is generalizable and replicable to documentary and information environments other than the health sciences.
- The project's focus is access and, by implication, preservation.
- Audiences for, and the value of, historical resources can be greatly expanded, and the relevance of such resources demonstrated through creative "packaging" or product development.
- Well-planned digitization offers endless possibilities for the manipulation and packaging of historical materials as user interest is perceived and information technology develops.
- Digital products must play an important role in cost-recovery and long-term self-sufficiency. The project would, therefore, prioritize inclusion of the most popular or most marketable historical resources held at each repository.
- An extensive collaborative approach is essential: archivists need to work with other disciplines and professions to make such a project work.
- Every participant has a unique and important perspective.
- Every participant must gain from the project, since MDHI is meeting multiple institutional, but not necessarily similar, goals and priorities.

After one and a half years of construction, the contents of MDHI include a statewide MARC-AMC-based guide to over 1500 archival collections nationwide relating to the history of the health sciences in Michigan; MARC catalogue records and associated images of medical and scientific artifacts held by the University of Michigan; imagebases of photographic and graphic art materials; technical manuals, protocols, and guidelines to assist communities engaged in similar digital projects; information about the location, collections, policies, and procedures of cooperating repositories; special documentary projects and exhibits such as one relating to the African American Health Care Experience in South East Michigan; information about publications and CD-ROMs available from SourceLINK; a digital version of the Historical Center for the Health Sciences's quarterly newsletter, *Retrospectives*, and links to related resources on the Internet.

As MDHI has grown, the limitations of the basic home page have become more evident. The limitations fall primarily into two categories: design and knowledge about use. In terms of design, one must bear in mind that most home pages using HTML are designed to be quick, inexpensive, and simple to develop, and therefore they lack the formality of a standard structural and conceptual framework such as that which could be generated using Standard Generalized Mark-up Language (SGML), as well as the sophisticated navigation and retrieval mechanisms provided by customized intelligent query interfaces or search engines. How this has played out with MDHI is that the maintenance of small files and escalating numbers of individually-created links is extremely labour intensive and structurally complex. In essence, it becomes increasingly difficult to get users quickly to the material they might wish to use, and even harder to get them back from that point to where they originated. The alternatives to a heavy reliance on hyperlinks for those who lack the

resources to move to a more sophisticated hypermedia development would be either more use of long linear documents (which requires considerable scrolling and does not exploit the power of hypermedia) or to use a hierarchy of menus rather similar to a gopher (which can also be laborious to navigate). Page design is also a problem since it can easily become inconsistent as new materials are added, HTML design capabilities continue to evolve, and navigational icons are revised.

In terms of use, like many other Websites, MDHI's user statistics quickly escalated, and currently run at more than sixteen thousand "hits" per month. While public domain software providing such quantitative assessments of home page use are readily available and can break user statistics down by broad categories such as place of origin, they are also able to provide home page owners with little more than public relations fodder. These statistics do not indicate, for example, how extensive are the uses underlying those "hits," who users are and what their purposes are in visiting the Website, and how many hits represent return visits.

MDHI plans to address these issues over the next year by developing a more comprehensive management and access structure that will support a context and content-rich digital library with a focus on bridging historical and contemporary information resources in the health sciences. MDHI is also looking at more detailed online quantitative and qualitative methods of tracking users and acquiring user feedback, such as logging the length of a visit, which files are accessed and how often, and perhaps even online user registration and exit interview procedures. Offline, MDHI is looking at the use of focus group interviews or think-aloud sessions to determine the relevance of MDHI content and the effectiveness of its design. As the first step in developing this new structure, MDHI is migrating in early 1996 to a new architecture called the Health and History Research Park (HHRP). HHRP will contain all existing MDHI resources, as well as several more, including a research facility to support and promote ongoing archival research into the design and use of digital libraries with historical content, an information centre for non-allopathic medical resources, and a virtual space for users to discuss their research among themselves, add their own annotations to materials they have viewed, or upload their own primary or secondary materials. HHRP will also be testing the effectiveness of the emerging SGML Document-Type Definition for Encoded Archival Description as the backbone structure for representing and relating archival description, contextualizing archival content, and linking to related digital objects.

The Potential of the Digital Library for Increasing Access to Health Sciences Documentation

The MDHI experience with a World Wide Web home page shows the potential for developing a comprehensive resource that addresses the vision outlined at the beginning of this article. It also demonstrates, however, the difficulties in continuing to expand in a hypermedia environment without adopting a coherent, scaleable, standards-based structure that will facilitate the development of a distributed network architecture; and without knowing more about audiences for such an environment. What then is the "digital library," and how might it resolve the shortcomings of home pages? The term "digital library" first came into currency around 1993, spurred on by the announcement of a joint funding initiative by the United States National

Science Foundation, Advanced Research Projects Agency, and National Aeronautics and Space Administration. This initiative was designed to bring academia, industry, and government together in digital libraries research and development. In doing so, the various agencies were recognizing that:

> To explore the full benefits of such digital libraries, the problem for research and development is not merely how to connect everyone and everything together in the network. Rather, it is to achieve an economically feasible capability to digitize massive corpora of extant and new information from heterogeneous and distributed sources; then store, search, process, and retrieve information from them in a user-friendly way.[6]

Indeed, an underlying premise of the digital library is the concept of "digital coherence." Once materials are in digital form, the "bits" from which they are composed can be related in all sorts of ways with the aid of document structures, intelligent agents, and computer algorithms to merge images, text, and even sound in ways that create whole new document forms and algorithmic information that can then be disseminated world-wide using global digital networks. Politicians and technologists predict that digital libraries are going to provide transparent accessibility to users across the multiple dimensions of time, space, economics, age, user groups, and physical capabilities.[7] How these digital libraries differ from home pages tends to be an issue of scale, resources employed, technological sophistication, economic viability, and level of conscious design.

Digital libraries with varying subject focuses are currently being developed with National Science Foundation funding at Stanford University, the University of California at Berkeley, the University of California at Santa Barbara, Carnegie-Mellon University, the University of Michigan, and the University of Illinois at Urbana-Champaign. Most of these digital libraries are being made available through World Wide Web interfaces, and incorporate all or some of the following characteristics:

- heterogeneous and distributed files and databases all linked together and cross-compilable through the use of intelligent agents;
- contents in multiple media (e.g., text, images, graphics, audio, video);
- universal accessibility from work, school, and the home through digital communications networks;
- designed for multiple and diverse user communities;
- all materials are electronic, most are digital (and therefore searchable), but some are analogue;
- dynamic—contents, nature, and uses are designed to change as the digital library and its users interact;
- economically feasible (e.g., cost recovery mechanisms such as chargeback or commercial components);
- frequently collaborative;
- hierarchies of both information and information surrogates;

- fewer materials available perhaps than in a traditional information setting, but fuller context developed for them;
- digital contents can be packaged, or other value added to them.[8]

Despite the considerable resources that are being invested by government, corporate, and academic institutions in the development of digital libraries, little historical content has been included to date. The most notable exception to this is the Library of Congress's National Digital Library Program (NDL). At the core of NDL is the American Memory Project which is including digital reproductions of archival materials relating to American culture and history, as well as finding aids and related materials.[9] When one looks at the potential for digital libraries with both historical and contemporary content for increasing access to documentation, however, health care and the health sciences seem to offer a particularly rich development opportunity. There is a demonstrable clinical and scientific need, for instance, to be able to look at data retrospectively, and also to correlate it with other information sources, and there have been discussions since the 1950s regarding the use of computers to integrate and manipulate health records for the purposes of conducting longitudinal demographic or cohort studies. At the same time, there are overarching ethical and privacy concerns and demands for effective cost models in the development of health information systems that map well onto current areas of research in digital library development. In addressing these demands and concerns, a digital library might contain algorithms for normalizing data elements between contemporary and historical sets of patient records; an online Institutional Review Board mechanism for reviewing requests to access those records; and even, using encryption or other secure transmission protocols, mechanisms for delivering complete or redacted records to approved researchers.

Howard Newcombe, an early proponent of medical networking, spoke with prescience in 1967 when he said that:

> For those who are interested primarily in the research uses of the medical and biological information contained in the various health records, it is necessary to remember that the custodians of the source files have as a primary responsibility either the production of routine statistics, the administration of schemes of health insurance and social welfare, or the maintenance of files for legal purposes. In these circumstances it is natural that acceptance of proposals for changing the existing methods of file upkeep so as to facilitate linkage on any substantial scale is likely to depend upon the demonstration of benefits relating either to the existing recognized uses of the files, or to essentially new uses such as the extraction of data in quantity for research purposes. In the latter case, however, the responsible agencies are apt to ask if the research products will be worth the additional cost.[10]

It is important, therefore, before launching into large-scale development of digital libraries in the health sciences, to identify and highlight areas where there is a commonality of interest with other stakeholders such as scientists, clinical researchers, or health policy makers. While this would be true of any potential development area, it is especially critical to be able to demonstrate the contemporary relevance of the history of the health sciences in the face of perennial cutbacks in funding for health care and the humanities, and where there is a desperate need for resources to be put

into addressing immediate healthcare needs. What, for example, might be the incentives for medical records administrators to get involved in digital library development? Could some of the functionality of a digital library also address key problem areas for contemporary medical records administration, such as more effective long-term tracking and compiling of outpatient, satellite clinic, or even shadow patient records, or ensuring that research records created under federal funding remain with the institution that received the funding and yet are accessible to the researcher who has moved to another institution?

Not all the uses of a digital library with history of health sciences content need be oriented toward the clinical and scientific user, however. The digital library could target a wider range of users, including the possibility of making value-added or packaged materials available to a large, emerging online user group: high school educators and students. In fact, one of the strengths of the digital library approach should be its ability to cater to heterogeneous user needs.

Functionality and Design of Digital Libraries in the Health Sciences with Historical Content

Bringing together a multiplicity of documentation types in order to present the user with a documentary whole in any subject area is an idea that has been discussed at length by Helen Samuels and Richard Cox. Most recently, Cox wrote that:

> Archivists talk about their mission to document society and then proceed to concentrate all their energies and resources on only one aspect of "documentation" ... Information valuable or essential to understanding any topic, geographic area, event, movement, individual's life, a family's development, or society can be found in a tremendous number of "sources." Artifacts, archaeological remains, popular culture, oral tradition, folklore, publications, movies and television, and archives and manuscripts are all essential for documenting society.[11]

Karen Sparck Jones, an information scientist approaching the issue of the different types of information contained in different types of information sources, has also underscored the value to the user of this more comprehensive approach to the information universe:

> We are dealing with [information] objects of very different kinds, not just physically as with pictures and text, but logically: pictures do not really have messages in the way texts do, and even with texts, for example a book, a news story, an acknowledgment signal, a purchase order, or a database entry, have very different sorts of messages. It is far from clear we can think of characterizing all of these in the same sort of way to access the information they contain ... We are meeting quite different sorts of relevance need, for example, extracting a personal record from a staff file to show a salary versus offering an abstract on a topic.[12]

Figures 1 and **2** illustrate the author's view of how this integration might look in terms of the contents and functionality of a digital library in the health sciences. **Figure 1** shows three content levels:

1. *Primary or base content*—This category represents the fundamental content of a digital library and pulls together a range of documentary materials and information contributed by contemporary record creators, as well as archives, museums, and libraries: reformatted archival material (e.g., digitized images and manuscripts); materials created digitally that could be inactive or archival in nature (e.g., electronic patient records or scientific research databases); data extracts (e.g., demographics extracted into a database from hospital admission records); materials created to redress documentary gaps (e.g., oral histories); digital descriptive information (e.g., repository guides, MARC records, and finding aids marked up using Encoded Archival Description [EAD]); bibliographic content (e.g., published research and secondary literature, indexes, abstracts, and bibliographies); and artifact content (e.g., images and descriptions of medical artifacts, instrument manufacturers trade catalogues).
2. *Value-added content*—This category speaks to some of the potential of digital libraries to be able to generate additional new content as well as revenue, and includes materials contributed by users (e.g., annotations of sources and personal research data); exhibits and educational materials that pre-package digital content in ways digital library creators believe will interest or be of use to targeted audiences; and data that have been extracted or cross-compiled from the library content according to user requests or commercial interest (e.g., demographics of Michigan sanitoria patients, 1925-1965).
3. *Externally available content*—This category exploits the capabilities of hypermedia to link to digital files of resources that exist outside the purview of the digital library. In general, although library developers have the power to create as many of these links as they see to be relevant, they do not have any control over the continued availability or quality of content of the materials to which they are linking. From a management and quality-control perspective, unless the digital library includes a mechanism for monitoring the sites to which it is linked, such links are likely to be most relevant when they are user, rather than system-defined.

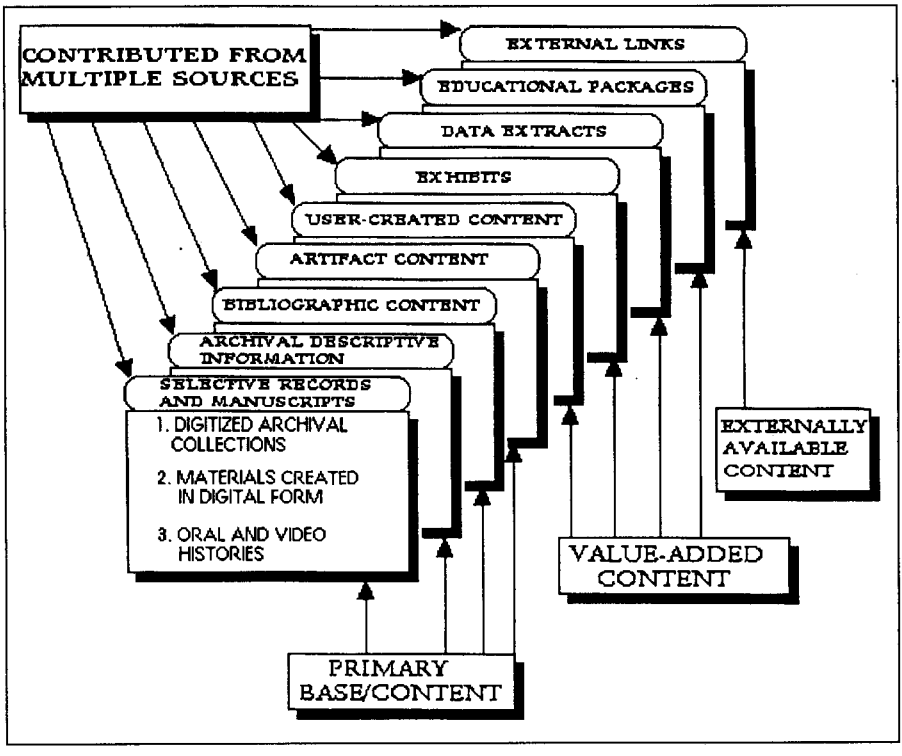

Figure 1: Potential Content of a Digital Library in the Health Sciences

Figure 2 indicates how a digital library might perform several complex technical functions (e.g., structuring and linking content, normalizing and mapping data elements, and compiling use information) at the same time as providing a variety of user services (e.g., different ways to browse and retrieve content, and online ordering and payment for images or value-added content).

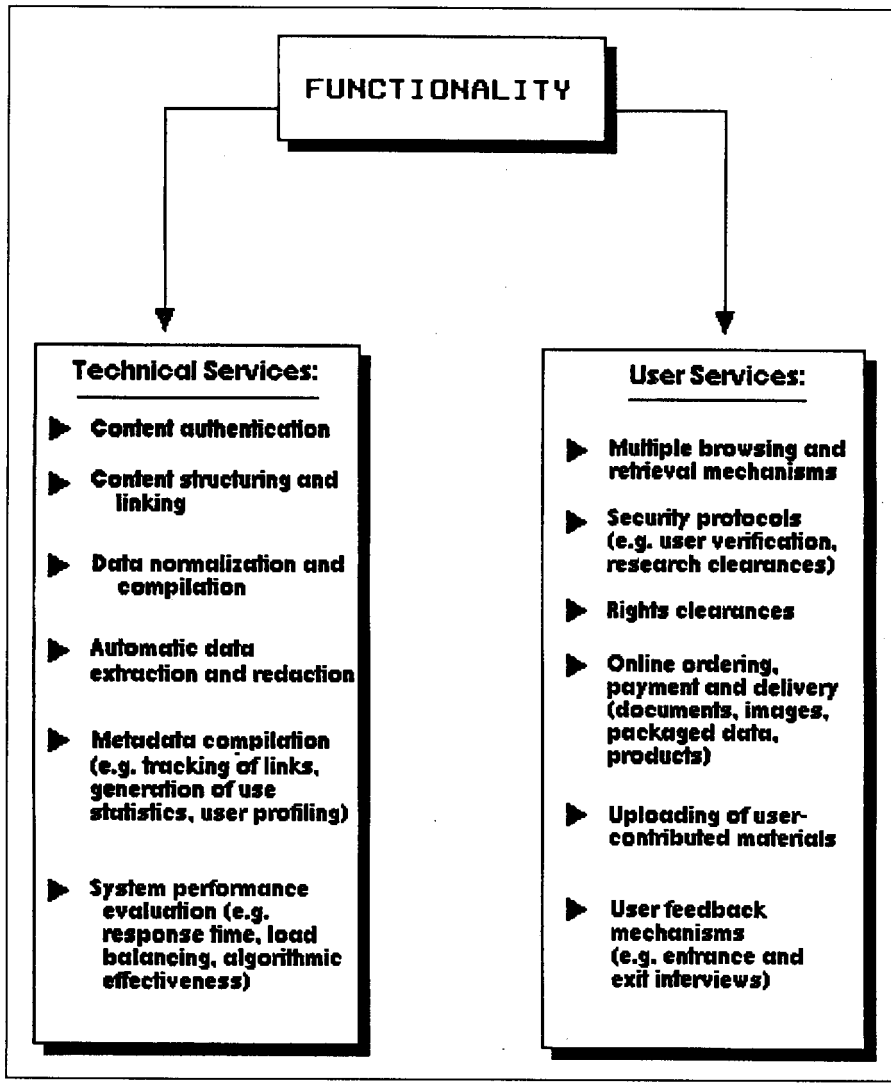

Figure 2: Potential Functionality of a Digital Library in the Health Sciences

Conclusion

In research areas such as the history of medicine and science, where examining all available evidence such as experimental data, researcher notes, published articles, and instrumentation are integral to deriving an understanding of clinical and research processes, digital libraries appear to facilitate both user access to, and synthesis of source material. For archivists in the health sciences and in general to become suc-

cessful participants in advanced networked hypermedia initiatives such as digital libraries, they are going to have to become increasingly access-driven in all of their activities, and be prepared to think outside traditional frameworks. If they do so, they give themselves unprecedented opportunities to promote their holdings and use them in new ways, reach out to new audiences, identify and redress documentary "gaps," share resources, and preserve physical collections. They also open up avenues to new funding and revenue-generating opportunities and the potential to enhance their own professional profile and role. The major stumbling blocks to achieving these benefits, however, are the tendency for archival digital programming to be driven by short-term exigencies rather than careful planning, systematic development, and an overarching vision, and the extent to which the archival profession is divorced from related activities that are occurring in other historical communities, and the health and information professions.

Notes

1 Karl Pearson, *The Grammar of Science* 3rd ed. (London, 1911).

2 This definition is taken, in part, from, *The Imaging Glossary of Electronic Document & Processing Terms, Acronyms and Concepts* (New York, 1991).

3 The scenarios were based on the following array of resources available to the author at the time of writing this paper: Anatomy: the Visible Human Project (http://www.nlm.nih.gov/extramural_research.dir/visible_human.html); the Digital Anatomist Program (http://www1.biostr.washington.edu/DigitalAnatomist.html); and the Art of Renaissance Science-Vesalius and the Study of Human Anatomy (http://bang.lanl.gov/video/stv/arshtml/renart2,html); cardiology: Preview the Heart, Heart History and Milestones in Cardiology (http://sin.fi.edu), American College of Cardiology (http://www.acc.org/), Cardiology Compass (http://osler.wustl.edu/~murphy/cardiology/compass.html), and a Cardiovascular Laboratory (http://www.vetmed.auburn.edu/~branch/CV1.html); microscopy: History of the Light Microscope (http://www.duke.edu/~tj/hist/huist_mic.html), Michigan Digital Historical Initiative, History of Science Museum, Florence, Italy (http://galileo.imss.firenze.it/index.html), Park Scientific Instrument (http://www.park.com.autosa.htm), Sutter Instrument Company Catalog (http://www.sutter.com/home.html); high school uses: Michigan Digital Historical Initiative and a Primer on Molecular genetics (http://www.gdb.org/Dan/DOE/introhtml); public health: Images from the History of the Public Health Service (http://www.nlm.nih.gov/hmd.dir/phs_history.dir/contents.html) and the "Virtual" Public Health Cente (http://www.sci.lib.uci.edu/HSG/PHealth.html); alternative medicine: Yahoo-Health, Alternative Medicine (http://www.yahoo.com/Health/Alternative_Medicine/), Foundation for Traditional Chinese Medicine (http://www.demon.co.uk/acupuncture/index.html), Complementary Medicine (http://galen.med.virginia.edu/~pjb3s/ComplementaryHomePage. html), and the Homeopathy Home Page (http://www.dungeon.com/~cam/homeo.html); genetics: Primer on Molecular Genetics, and Amish Genetic Diseases (http://www.crynwr.com/amish/).

4 Clifford Stoll, *Silicon Snake Oil: Second Thoughts on the Information Highway* (New York, 1995), p. 2.

5 For further information on SourceLINK and MDHI, see Anne J. Gilliland-Swetland, ed., *The History of the Health Sciences in Michigan: A Guide to Research Sources 1994-1995* (Ann Arbor, 1995) and "CD-ROM Development at the University of Michigan Historical Center for the Health Sciences," *Watermark* XVII, no.1 (Winter 1994), pp. 9-11; Denise Anthony, "The Michigan Digital Historical Initiative," *Watermark* XVIII, no. 3 (Summer 1995); and World Wide Web URL: http://http2.sils.umich.edu/HCHS/.

6 *Research on Digital Libraries, Announcement* (National Science Foundation Computer and Information Science and Engineering Directorate, Advanced Research Projects Agency Computing Systems Technology Office, and the Software and Intelligent Systems Technology Office, National Aeronautics and Space Administration, 1993).

7 See also David M. Levy and Catherine C. Marshall, "Going Digital: A Look at the Assumptions Underlying Digital Libraries," *Communications of the ACM* 38 (April 1995), pp. 77-84.

8 For a fuller description, see Anne J. Gilliland-Swetland, "Computer-based Communications and Archives: Documentary Opportunities Not to be Missed," *Archival Issues* (Fall 1995), forthcoming.

9 See Carl Fleischauer, "Elements of Digital Archival Collections: Technical Overview and Format Description," (27 October 1994); and "Organizing Digital Archival Collections: American Memory's Experiences with Bibliographic Records and Other Finding Aids," (n.d.). Both papers are available from the Library of Congress Website.

10 Howard B. Newcombe, "Products from the Early Stages of Development of a System of Linked Records," in E.D. Acheson, ed., *Record Linkage in Medicine, Proceedings of the International Symposium, Exford, July 1967* (Baltimore, 1968), p. 7.

11 See Richard Cox, "The Documentation Strategy and Archival Appraisal Principles: A Different Perspective," *Archivaria* 38 (Fall 1994), pp. 11-37.

12 Karen Sparck Jones, "Fashionable Trends and Feasible Strategies in Information Management," *Information Processing and Management* 34, no. 6 (1988), p. 706.