RICHARD E. LUCIER

University Librarian and Assistant Vice Chancellor
for Academic Information Management
University of California, San Francisco

Embedding the Library into Scientific and Scholarly Communication through Knowledge Management

ABSTRACT

Knowledge management is a new role for academic research libraries that has the potential to integrate the library into scholarly and scientific communication in a significant way. Work in knowledge management is advancing in both the sciences and humanities. The Genome Data Base at the Johns Hopkins University is currently the most advanced knowledge management prototype. As part of its new Center for Knowledge Management, the University of California, San Francisco is undertaking several initiatives to create a campuswide knowledge management environment.

INTRODUCTION

The University of California, San Francisco (UCSF) is one of the nine campuses of the University of California (UC) system. With schools in medicine, pharmacy, nursing, and dentistry, and graduate programs in the behavioral and social sciences, UCSF is unique within UC in that it is the only campus devoted to research, education, and service in the health sciences.

In September 1990, as the result of a decade of planning, UCSF opened a new library building of great beauty and utility that is a visual representation of the importance of the library to the UCSF faculty and student community (Cooper, 1991). The critical challenge in the current decade is to articulate and realize a programmatic vision that will (a) embed the library into the scientific and clinical research, educational curricula, and professional practice programs of this diverse and distributed campus; (b) position the library as a campus focal point for knowledge-based applications of information technology; and (c) establish the library's leadership in the development of knowledge bases and online tools for the health sciences.

OUR VISION: A KNOWLEDGE MANAGEMENT ENVIRONMENT

Historically, the function of the research library has been storage and retrieval. This will remain at the core of the library's responsibilities. More recently, the library has extended its role to include information transfer, or the delivery of information over high-speed communications networks. Responsibilities and activities in this area are increasing rapidly, driven by users' needs and the growing availability and reliability of the Internet or the National Research and Education Network (NREN). A new, more experimental and challenging role for the library is that of knowledge management, the insinuation of the library at the beginning of the scientific and scholarly communication process for the purpose of building and maintaining specialized knowledge bases in unique collaborations with scientists and scholars.

Our vision for the UCSF Library, and its innovative new Center for Knowledge Management, embraces all three functions: storage and retrieval, information transfer, and knowledge management. Figure 1 graphically represents this vision, which we call a Knowledge Management Environment.

This Knowledge Management Environment is an integration of knowledge sources, access and delivery systems, education and training programs, and personalized services with the following components:

- online bibliographic databases of the library's physical collection;
- the "full text" of the published literature online, including images;
- high-quality, interactive knowledge bases *critical* to the daily work of scholars and scientists;
- online tools for the peer review of data and collaborative knowledge base management;
- high-speed communications for the conduct of scientific and scholarly work from the local to international levels; and

- an integrated access tool, or wide area information server, to retrieve information from local and remote bibliographic databases, "full-text" information sources, and specialized knowledge bases.

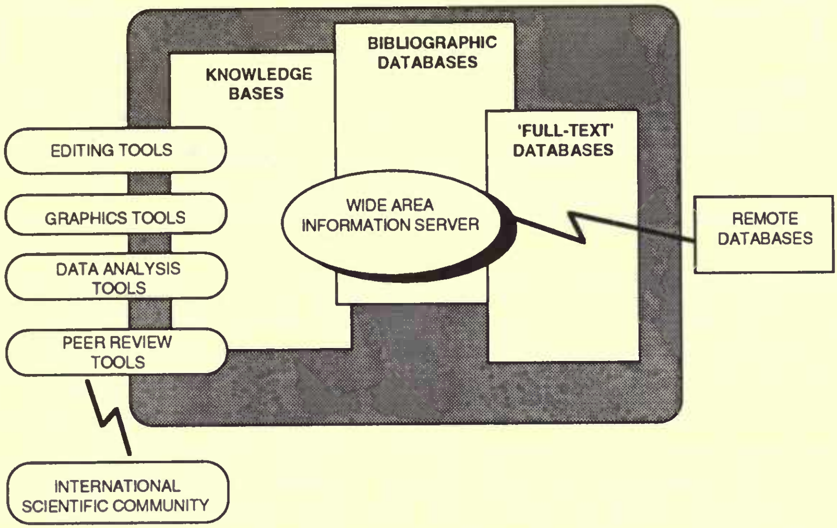

Figure 1. The Knowledge Management Environment

SCIENTIFIC AND SCHOLARLY COMMUNICATION

The need for a Knowledge Management Environment emerges from problems inherent in the current scientific and scholarly communication process. Figure 2 depicts the information transfer cycle as we know it today. Scientists and scholars discover new knowledge and communicate it through both writing and teaching. Publishers disseminate that information through a variety of primary and secondary information products. In their traditional storage and retrieval role, research libraries build collections and make available to users the world's published literature. Since the 1970s, network access (information transfer) to this stored knowledge through online catalogs and indexes, along with a new emphasis on service and education, has assumed major importance. However, the roles and functions of scholars, publishers, and librarians have remained fundamentally the same.

By the late 1980s, the limitations of this prevailing model for scientific communication were becoming apparent. The length of the hard-copy publishing process makes it increasingly difficult for scholars

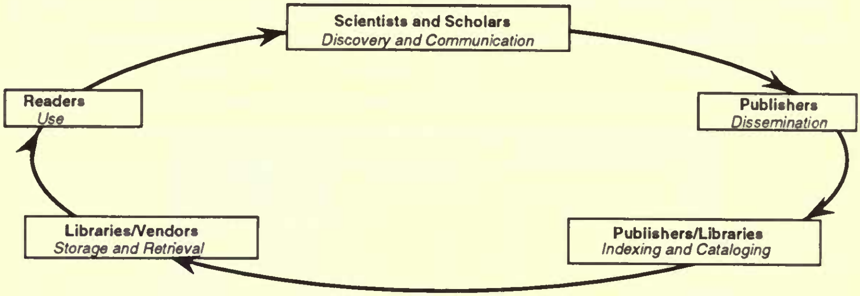

Figure 2. The information transfer cycle

and scientists to communicate their findings in a timely fashion. With the rising cost of publishing and a limited resource base, libraries and universities can no longer afford to support comprehensive collections. The financial crisis facing libraries is not short-term; rather, it is structural in the current environment. Most importantly, it is clear that the presentation of knowledge in static form, whether in print or as part of the emerging electronic library, is grossly inadequate. Scientists and scholars, often on their own and with inadequate support, are augmenting this passive presentation of knowledge with a growing number of interactive, discipline-based knowledge bases that are developed, maintained, and shared across networks. Knowledge management has emerged from this situation as a creative response to managing the world's knowledge base.

THE LIBRARY AS KNOWLEDGE MANAGER

Knowledge management represents a new model for scientific and scholarly communication in which faculty and research librarians share the responsibility for the collection, structuring, representation, dissemination, and use of knowledge using electronic information technologies. Encompassing the entire information life cycle, from creation of new knowledge to its dissemination and use, knowledge management is a collaborative enterprise, where scholars, scientists, and research librarians work together to develop and maintain knowledge bases and derivative information products. Knowledge bases are developed and maintained through knowledge management processes, which ensure content integrity and usefulness. A variety of products and services can be derived from the knowledge base. The collaborative nature of knowledge management, embodied in techniques of shared development of functional specifications, rapid prototyping, and user acceptance testing, fosters an interdependency among all involved. The ongoing management of the knowledge base also requires funding and administration strategies that crosscut traditional departmental, disciplinary, and institutional boundaries.

Knowledge management consists of four primary components:

1. Collaboration: the shared responsibility for the development and management of knowledge bases, products, and services. Effective collaboration requires a balanced relationship among peers, recognizing the unique value of each person's contributions to the success of shared work. A multidisciplinary team of collaborators includes discipline-based scholars and scientists, librarians, computer scientists, and software engineers.
2. Knowledge base: a collection of scholarly knowledge structured for computational storage and representation. A knowledge base may contain all or some part of the intellectual core of a scholarly discipline. The contents of the knowledge base are chosen and validated by consensus at some level within the scholarly community that develops, uses, and maintains it.
3. Knowledge management processes: those activities of collaborators related to the creation, structuring, representation, dissemination, and use of scholarly knowledge. They result in knowledge bases, patterns of collaboration and communication that ensure the integrity and continuing usefulness of those knowledge bases, and knowledge products.
4. Knowledge products and services are the output derived from the knowledge base: books, articles, computer-based educational materials, database subsets, and typesetting tapes are examples of knowledge products. Knowledge products are market driven, developed in response to the immediate information needs of scholars, scientists, educators, students, and other information seekers. Product services are the customer support activities associated with each knowledge product. Examples include production of typesetting tapes or camera-ready copy for hard-copy publication or education and training programs to provide skills and abilities needed for full and appropriate use of knowledge products derived from the knowledge base.

What is remarkably different about the knowledge management role is that it insinuates the library at the beginning of the information transfer cycle rather than at the end and focuses on information capture rather than access and use (Figure 3).

The long-term implications for building and maintaining a portion of the library's collection in this manner are enormous. Knowledge management transforms the various roles in the scientific communication process and potentially places ownership and control back in the hands of the scholarly community. It also has enormous potential for closing the gap between research faculty and their students and integrating the library into research and education programs in a significant way.

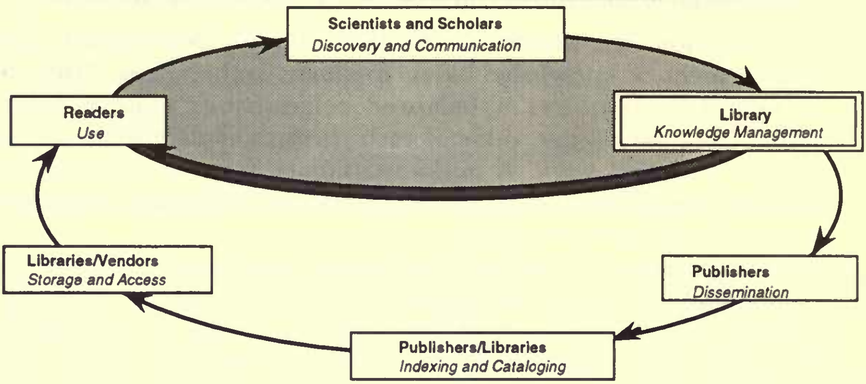

Figure 3. New roles in a Knowledge Management Environment

THE GENOME DATA BASE: A KNOWLEDGE MANAGEMENT PROTOTYPE

To date, the most advanced knowledge management prototype is the Genome Data Base (GDB), developed at the Laboratory for Applied Research in Academic Information, William H. Welch Medical Library, the Johns Hopkins University. (The following section is an adaptation of sections from Lucier [1990].) GDB is a working prototype, which serves the international scientific community on a daily basis. The most technologically advanced systems possible are not our primary goal in knowledge management; instead we are more concerned with designing systems that work and that people use in their everyday environments.

GDB is a gene-mapping database that serves as a repository for data collected by scientists engaged in the international human genome effort. GDB integrates several types of data including descriptions and map locations of human genes and other markers, descriptions of DNA probes used to characterize the markers and polymorphisms, contacts for obtaining probes, and more than 25,000 linked bibliographic citations.

To see GDB as an example of knowledge management, it is essential to have an understanding of the sociology of the human genetics community, namely the Human Gene Mapping (HGM) Workshops. The First International Human Gene Mapping Workshop, held in 1973, was instituted to develop and maintain a consensus human gene map. Since that time, similar workshops have been held either annually or biennially. These workshops are one of the community's primary data filters. The HGM workshops are organized by committee, one for each chromosome as well as several specialized committees, e.g., nomenclature, DNA, mitochondrial, and comparative (mouse). These committees collect, review, analyze, and synthesize all the mapping data from the published literature to produce the consensus human gene map.

Two aspects of HGM work have specially driven the need and design for GDB: (a) the growing volume and complexity of data and (b) the interactive character of the peer review work of HGM committees. The amount of information that committees must process has increased proportionately with the greatly heightened scientific activity in this area. It is estimated that the information doubles every two years. In 1973, 75 people attended HGM 1, and 25 genes were mapped. At HGM 10 in 1989, 700 scientists were in attendance, and 1,630 genes were mapped. Until now, during the four- or five-day workshop, members would collect and input information concerning their particular chromosome. With the larger volume of data, this has become a nearly impossible task, even with the introduction of computers to the workshops beginning in 1983. Making use of the Internet and public data networks, GDB provides the committees with an online, continuous, interactive system into which information can be added and verified at the committee members' convenience throughout the year. The committees will continue to meet annually, but the workshops can now focus on science and the analysis and significance of data rather than on data entry.

Although the various chromosome committees do their work mostly independent of each other, there is considerable interaction among the chromosome, nomenclature, and DNA committees. Certain data elements are shared; these elements, e.g., gene symbol, cannot become part of the database until they have been validated by the appropriate members of various committees, in consultation with each other. An "online peer review process" has been integrated into GDB editorial interfaces, again making significant use of national and international networks and a completely modularized design.

At present, GDB draws primarily on the HGM workshops and the literature for the major portion of its data. Already, it is beginning to include unpublished and unvalidated data submitted directly by users for consideration and subjected to quality control by both GDB staff and a special group of scientific editors. Figure 4 illustrates this data flow and highly dynamic form of scientific communication possible in a networked environment. It also represents a true electronic journal in a knowledge management environment.

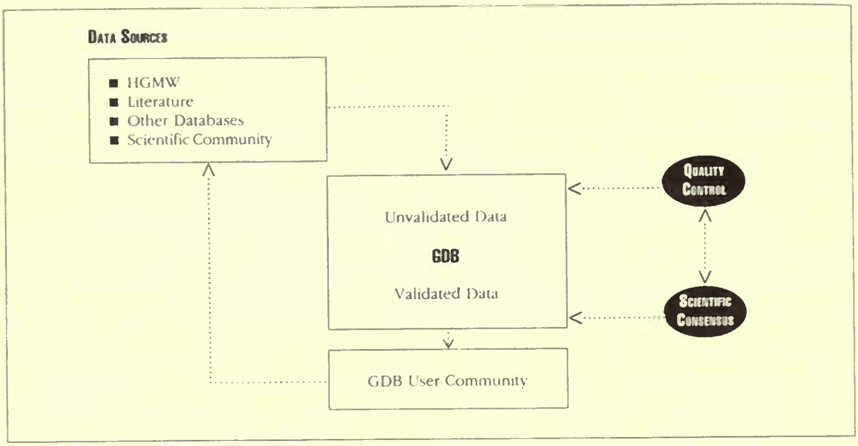

Figure 4. GDB data flow in a networked environment

GDB is designed so that it is possible to develop other information products that the user community demands in order to accomplish its work. In addition to the various interfaces provided for HGM committees and GDB Editors, a more generalized online searchable version of GDB is available to the scientific community. The HGM Reports, published by Karger in a special issue of the journal *Cell Genetics and Cytogenetics*, are produced from GDB data structures.

IMPLEMENTATION OF A KNOWLEDGE MANAGEMENT ENVIRONMENT AT UCSF

Although the library will continue to build a high-quality paper-based collection in the health sciences, excellent service in a distributed environment as well as educational programs will assume a far higher priority than in the past. We anticipate a rapidly increasing emphasis on information transfer and knowledge management over the next 10 years, and we will focus our technology-based efforts on these roles.

Figure 5 depicts the primary areas in which we plan to develop or adopt technological innovations over the next three to five years, as we implement the first phase of our Knowledge Management Environment.

Driven by the needs of our customers, the continuously changing external environment, and new advances in technology, we are fashioning a dynamic, multidisciplinary organization with three programmatic divisions.

	CONTENT	ACCESS	EDUCATION	SERVICE
STORAGE & RETRIEVAL	Bibliographic db's Online Indices 'Full Text'			
INFORMATION TRANSFER	Delivery of 'Full Text' information on-demand	Wide Area Information Server (GALEN) for integrated access to local/remote db's	Training programs which facilitate access and use	Information services to distributed faculty and student workplaces
KNOWLEDGE MANAGEMENT	Discipline-specific knowledge bases of high currency, value, and integrity Interactive tools for collaborative data maintenance in a networked environment	Information Retrieval Software for 'Full Text' Electronic Library Online tools for information access and analysis	Practice-based health sciences Informatics Curriculum Information products to support curriculum	Direct collaboration between faculty and librarians in the development/ maintenance of knowledge bases and products

Figure 5. Development areas for the first phase of the Knowledge Management Environment

1. *Information Resources and Services* is responsible for storage and retrieval and information transfer functions including the following:
   - collection management and processing;
   - public information services to our distributed customer base;
   - document and information delivery;
   - special collections including the History of the Health Sciences, the Oriental Collection, and University Archives; and
   - bibliographic instruction component of a broad educational program.
2. The *Center for Knowledge Management*, created by the Division of Academic Affairs and the library, is responsible for information transfer and knowledge management functions. In collaboration with faculty, the center's staff performs the following functions:
   - develops new information products and services, e.g., knowledge bases and online tools for the health sciences;
   - pursues applied research projects related to UCSF informatics problems;
   - advises graduate students in computer science and the health sciences who are using the center as their laboratory;
   - consults with faculty, staff, and students in the development of private databases, etc.; and

- supports the state-of-the-art systems and infrastructure that underpin the development, maintenance, and use of knowledge resources and information services.

3. The *Interactive Learning Laboratory* has primary responsibility for our educational and instructional programs including the following:
   - development of a health sciences informatics curriculum;
   - integration of educational technology resources into the curriculum of the various schools and professional training programs;
   - instructional computing and the development of multimedia software for education; and
   - educational and external publications.

A fourth division, *Finance, Planning, and Administration,* supports our storage and retrieval, information transfer, and knowledge management functions through the efficient and effective management of our financial and human resources and facilities. This division is also responsible for development. In order to implement the UCSF Knowledge Management Environment, it is critical for the library to implement long-term financial planning for the effective use of state funds as well as broaden its financial resource base beyond state-appropriated funds. Important sources of support include grants, contracts, business-university agreements, gifts, and information consulting and brokering activities. An endowment for the Center for Knowledge Management has been established as an important priority in an upcoming campus capital campaign.

As we move towards realizing our Knowledge Management Environment vision, it has also been necessary for us to examine and refashion the library's organizational culture as well. Several principles guide us in this challenging and long-term task that has been greatly aided by the opportunity to recruit several new professional staff from the library, computing, and biomedical science professions:

- high value placed on technological innovations that solve practical and recognized problems;
- continuous involvement of faculty, staff, and students in the University of California tradition of shared governance;
- an informed, knowledgeable, and service-oriented staff a critical factor;
- technology a tool, not an end;
- strong management essential for program development and the effective use of human and financial resources;
- processes and tasks organized around outcomes;
- pragmatism and principle as a dual basis for decision making;
- outcomes as the principal evaluation measure; and

- entrepreneurial responsiveness to environmental changes, opportunities, and emerging information technologies a key to success.

Several new projects have already begun. In collaboration with our Human Gene Mapping Center, we have successfully sought funding to build and maintain a chromosome 4 database, which will be our first efforts at collecting and making available source data. Discussions are continuing with (a) Springer-Verlag for an experiment with several important online journals, (b) clinical researchers for an AIDS knowledge base, and (c) medical educators for the creation of a comprehensive database that would support undergraduate medical education.

CLR STUDY OF KNOWLEDGE MANAGEMENT

In 1987, the Council on Library Resources (CLR) awarded a grant to Richard E. Lucier and Nina W. Matheson to address the changing roles of research libraries, the scholarly community, and university publishers in scientific and scholarly communication through examination of the knowledge management model as implemented in Lucier's work at the Laboratory for Applied Research in Academic Information, the William H. Welch Medical Library, the Johns Hopkins University. (This following section is an adaptation of sections from Lucier & Matheson [1992].) The CLR grant had three major objectives:

1. *Documentation of the knowledge management model.* The collection, examination, and synthesis of statements, definitions, and descriptions of the knowledge management model and its components have been major documentation activities of the project. Briefing materials for the Symposium on Knowledge Management drew heavily upon these files. A monograph on the knowledge management model, coauthored by the principal investigators, will be published by the Johns Hopkins University Press in late 1992.
2. *Diffusion of the knowledge management model to academic settings outside medicine.* Initial diffusion of the concept of the model occurred through presentations made by the principal investigators to high-level staff at numerous academic institutions that seemed to possess the requisite human, technical, and financial resources to implement the model. Follow-up calls, interviews, and site visits monitored the possibility of actual implementation in these settings. Presentations were also made at several national meetings over the life of the grant.
3. *Sponsorship of a national meeting on knowledge management.* Early on, the principal investigators formed a special executive committee

to oversee this component of the project. This group decided on the strategy of a special invitational symposium as most appropriate to a full discussion of the key issues raised by the knowledge management model.

From October 27-29, 1991, 63 invited guests gathered at the Coolfont Conference Center in Berkeley Springs, West Virginia, for the Invitational Symposium on Knowledge Management, a policy-level forum for examination of knowledge management. Included among these experts were scholars, university administrators, academic librarians from major public and private universities, association directors, independent consultants, and others whose work and interests have led to innovations in scholarly and scientific information management. Major private and public funding agencies such as the National Science Foundation, the National Endowment for the Humanities, and the Andrew W. Mellon Foundation were also represented. During the three-day gathering, participants' time was divided between working group meetings and plenary sessions.

Symposium registrants directed their attention and activities toward four desired outcomes:

1. *Shared understanding* of the knowledge management model, including the economic and political advantages and disadvantages of different approaches and social and other noneconomic barriers to wider implementation of knowledge management.
2. *Clarification of implications* for scientific/scholarly communication, comparing the current situation to communication in knowledge management environments, and suggesting solutions for problem areas.
3. *Scenario development* outside human genetics, applying knowledge management to other scholarly information problems; examining existing knowledge management projects, especially in the humanities; and identifying the advantages, disadvantages, opportunities, and barriers to knowledge management within particular disciplines.
4. *Recommending implementation strategies* for knowledge management, providing a rationale, time frame, level of intensity, projected resource requirements, technological initiatives, and, where possible, priority audiences.

At the symposium's concluding session, the leaders of these groups presented recommendations for future actions in each of the five areas. *Strong consensus emerged in support of wider implementation of knowledge management.* In the words of Donald N. Langenberg, registrants should take action to colonize carefully selected distant locations in intellectual space with the practice of knowledge management. Groups also pinpointed a set of complementary actions to develop the conceptual and curricular infrastructure for knowledge management environments.

Knowledge management is a transformational activity. Working examples of knowledge management serve as proof of concept for the approach. They also help to highlight the areas where immediate work is needed if an infrastructure to nurture new implementations is to emerge in the next three to five years. Next steps involve actions with national and international impact; individual initiatives must be supplanted by broader based, mainstream action targeted to reduce barriers and leverage opportunities.

Themes running through the plenary and working group discussions and recommendations highlight three action items that require immediate attention:

1. *Financial strategies.* The future diffusion and success of knowledge management rests in large part on the development of reasonable and creative financial strategies and on an economic model that considers the needs of all important players. In particular, the model must consider that the current state of research institutions and information producers, in an era of limited resources and constrained public agencies, requires the ability to mesh pricing, costing, and allocation strategies among various organizations and groups both internal and external to the research enterprise.
2. *Intellectual property.* As a next logical step in moving towards a knowledge management environment, it is increasingly important to convert existing published works to electronic form for online access and management. The symposium's work group on intellectual property proposed pulling a group together to describe the climate needed for knowledge management, particularly the elements of collaborative ownership. Such a group would identify current copyright status for each class of information (e.g., source data, consensus data, the published literature, and bibliographic records), project what is needed, and recommend actions to be taken over the next five years.
3. *Technology strategy.* A functional architecture that will serve as a reference model is needed for knowledge management. This structural definition can serve as a rationale for institutional infrastructure planning and technological development. The architecture should take into account the available technologies but must also offer a plan for incorporating future developments. Though there will ultimately be several architectures for knowledge management, a general one is needed to begin with that defines how to deal with communications, content organization, standards, and other related issues.

CONCLUSION

It is clear that the knowledge management concept is a vital, effective approach to scientific communication in networked environments. Working implementations of knowledge management exist, and numerous projects in the sciences and humanities can be identified where the knowledge management approach will provide identifiable benefits to disciplines and institutions. Wider implementation of knowledge management approaches requires that the focus of action and attention be redirected to issues beyond those that arise from individual university- or discipline-specific projects. Enthusiasm exists for initiating new knowledge management experiments in a number of disciplines, but it is not likely that any coordinated effort can emerge until additional work is done to reduce technological, legal, and financial barriers. The involvement of new participants, including people who bring legal and economic expertise and who share an interest in and commitment to shaping new roles and processes in scholarly and scientific communication, is critical.

REFERENCES

Cooper, R. S. (1991). A library for the fifteenth through the twenty-first centuries. *Bulletin of the Medical Library Association, 79*(2), 147-158.

Lucier, R. E. (1990). Knowledge management: Refining roles in scientific communication. *EDUCOM Review, 25*(3), 21-27.

Lucier, R. E., & Matheson, N. W. (1992). *Invitational symposium on knowledge management: Overview and recommendations.* Baltimore, MD: Welch Medical Library.