The Virtual Library:
What Is It and Where Are We Headed?

by Gary Harden, Guest Editor

Much has been written recently about the "virtual library." Some say that the virtual library will make librarians obsolete. Others say that the library profession is on the verge of a major transition and stands to gain in professional stature from the development of virtual library services. So, what is the virtual library and what are the potential effects on library service and librarianship?

To understand how the virtual library concept has developed, a bit of historical background is in order. The foundation upon which the concept is based is the development of high-speed computing and networking technology. Although the development of computing machinery had its beginnings with Charles Babbage's difference engine and Herman Hollerith's tabulating machine in the 1800s,[1] the modern era of digital computing began in the 1930s. In 1936 Howard Aiken, a Harvard professor of mathematics, approached Thomas Watson, Sr., of IBM with a proposal to build a large scale computing machine. Watson was impressed by the design and provided $1 million in funding for the project. The resulting Harvard Mark I was unveiled in 1944.[2] Around the same time Dr. John Atanasoff, a professor of physics at Iowa State University, designed the first true electronic digital computer. As the United States entered World War II, the military became interested in machines that could be used in the war effort. Dr. John Mauchly of the University of Pennsylvania and his student assistant, J. Presper Eckert, began work on a computer for the military that could calculate artillery and missle trajectories. Their design was a refinement of Atanasoff's digital computer. The ENIAC

(Electronic Numerical Integrator and Calculator) was completed in February 1946, too late to be used in the war effort.[3] The ENIAC was cumbersome to use because its programming was "hard-wired." Each time that a new program was to be run, the machine physically had to be rewired. Dr. John Von Neumann, a mathematician and team member of the Manhattan Project, proposed a new design which would use the stored program concept. The EDVAC, or Electronic Discrete Variable Automatic Computer, would store all program instructions in computer memory. Switching from one program to another could now be accomplished quickly and easily. This set the stage for the beginning of the Information Age.[4]

In the relatively short span of forty years, computing technology has developed to an astonishing degree. The ENIAC occupied fifteen hundred square feet of space and weighed thirty tons. The microcomputer of today fits on a desktop and is many times more powerful than the original ENIAC. This reduction in size and increase in power is one factor leading to the development of virtual library services. The other major factor in this development is the maturation of networking technology. Effective information delivery depends upon the ability to interconnect disparate computing platforms into a cohesive network that utilizes a standard communications protocol. A network which could provide this functionality began to take shape in 1969. The ARPANET was established by the Department of Defense Advanced Research Projects Agency (which became known as DARPA) as an experimental network linking researchers at academic institutions and government laboratories. Over time other networks were interconnected with ARPANET using the commu-

nications protocols (TCP/IP) developed by DARPA. By the early 1980s, the ARPANET had split into two interconnected networks and the "Internet" was born. The single most important factor contributing to the explosive growth of networking was the establishment of the National Science Foundation Network (NSFNET) in 1986. From this point forward, the Internet has expanded to become an open global network interconnecting thousands of local and regional networks.[5]

Research and development during the last half-century have given us the means to implement the "virtual library," but what exactly is it? The virtual library can be defined as a collection of books, documents, images, recordings, etc., that is stored in binary (machine-readable) format and which can be accessed through electronic means. The concept implies that one can have access to the contents of materials without having the physical materials themselves. Dr. Vannevar Bush, Director of the Office of Scientific Research and Development during World War II, hinted at the virtual library concept in a landmark article published in *The Atlantic Monthly* in 1945. Bush recognized the overwhelming growth in humankind's vast store of knowledge and the difficulties inherent in accessing specific information: "The summation of human experience is being expanded at a prodigious rate, and the means we use for threading through the consequent maze to the momentarily important item is the same as was used in the days of square-rigged ships."[6] The visionary Bush foresaw the development of sophisticated computing devices that would facilitate the organization and distribution of information: "The advanced arithmetical machines of the future will be electrical in nature, and

they will perform at 100 times present speeds, or more. Moreover, they will be far more versatile than present commercial machines, so that they may readily be adapted for a wide variety of operations."[7] Bush believed that specific data became increasingly difficult to locate as the volume of stored information grew. He felt that the hierarchical systems of indexing in use at the time were artificial and did not reflect the natural processes of the human mind: "The human mind ... operates by association. With one item in its grasp, it snaps instantly to the next that is suggested by the association of thoughts ... Selection by association, rather than by indexing, may yet be mechanized."[8] Bush visualized a device for personal use which could function as a mechanical file and library. He called this device a "memex." The memex would store books, records, communications, etc. and would enable this stored information to be searched quickly and in a flexible manner. He wrote, "It affords an immediate step, however, to associative indexing, the basic idea of which is a provision whereby any item may be caused at will to select immediately and automatically another. This is the essential feature of the memex. The process of tying two items together is the important thing."[9]

The desktop microcomputer, with communications links to the global Internet, can be viewed as the logical extension of the memex conceived by Vannevar Bush. Microcomputers have developed to such a degree that they can be used to retrieve, store, and manipulate prodigious amounts of data in every conceivable format: text, graphical images, video files, audio files, binary files, etc. Through the use of graphical client software (such as Mosaic) to access servers running HTTP (Hypertext Transfer Protocol), the concept of "associative indexing" moves closer to becoming a reality. HTTP provides the means to create "linked" documents across many different computer systems. The researcher can follow a specific information "thread" across the Internet by choosing associated links in HTTP documents. These links can point to related documents in different formats. For example, one could connect to a server offering a biographical text file on Doc Watson, choose a link in the document which would connect to a document containing photographs of Watson and his family, and then choose a link which would connect to audio files of an actual performance by the famous musician. Different types of information in different formats brought together through association — the essence of the memex.

A number of electronic library projects

have developed over the past decade. Notable among these projects are the Mann Library Gateway at Cornell University, Project JANUS at the Columbia University Law School, and the Library of Congress' American Memory Project.

The Albert R. Mann Library at Cornell University implemented an electronic library service in 1991. The service, known as the Mann Library Gateway, includes numeric data files, computerized journal indexes, corporate profiles, library catalogs, and the full text of journal articles and reference works.[10] This electronic library project was developed by the Public Services, Collection Development, and Technical Services divisions in cooperation with the Information Technology Section. The Mann Library received the first ALA / Meckler Library of the Future award in 1993 for their work on the Mann Library Gateway.[11] Access to the electronic library is currently restricted to the Cornell University community due to database restrictions, but a "public" version of the Gateway will eventually be available to anyone on the Internet.

In 1990, the Columbia University Law School Library was charged with developing an alternative mode of library access which would preclude the expansion of existing space. The law librarian, James Hoover, and the director of Computer Systems and Research, Willem Scholten, developed a virtual library concept based on the use of a supercomputer. Project JANUS took shape in November 1992, when a CM-2 supercomputer from Thinking Machines Corporation was installed in the Law Library. This is the first library on record to utilize a supercomputer for virtual library services.[12] "Project JANUS is a prototype digital library which utilizes the power of a massively parallel supercomputer to provide users with access to texts, images, sound, and video from remote and local workstations. As the JANUS project is developed, users will be able to have access to tens of thousands of books, both archival and current copyrighted editions. In addition, JANUS is a means of preservation and enhanced access to archival collections such as the Perlin Papers (the Rosenberg/Sobell FBI Surveillance Archive) and the Nuremburg Trial Papers."[13]

The American Memory Project began development in 1990 as a means to provide electronic access to the Library of Congress' collections of archival materials. A variety of multimedia materials are available, including films, audio recordings, broadsides, and photographs. As initially conceived, American Memory was provided on videodisc to forty-four libraries.[14] Recently the Library of Con-

gress implemented an HTTP (World Wide Web) server which provides Internet access to the American Memory collections. (http://marvel.loc.gov/homepage/lchp.htnl) These valuable historical collections are now accessible to any library having an Internet connection.

The virtual library is no longer just a concept. As the cited examples demonstrate, it exists now in various forms at numerous institutions around the globe. Although still in its developmental stages, the virtual library of today portends the future library of tomorrow, but technological developments in any field are not without consequences. Are physical libraries and librarians becoming obsolete? Are books in danger of disappearing? These questions currently are being debated among library professionals. Michael Gorman writes, "Libraries are under attack as never before, and none more so than academic libraries. The enemies of academic libraries fall into three classes. The *bureaucrats* know little or nothing of education or libraries. They know only that they cost a lot of money; money that could be saved if libraries were to be dismantled behind a smokescreen of technology. The *technocrats*, or at least some of them, believe that technology can be used to provide something equal to, or better than, 'traditional' library services. The *technovandals* want to use technology to break up the culture of learning and, in a weird mixture of nineties cybervision and sixties radicalism, to replace that world with a howling wilderness of unstructured, unrelated gobbets of 'information' and random images in which the hapless individual wanders without direction or sense of value."[15]

Mr. Gorman argues (rightly, I think) that the book and libraries must be saved from destruction. On the other hand, he assumes that the book can take only one form, that of ink on paper. John Kountz observes that "In the next five years or so, the market for — and the availability of — information printed on paper can be anticipated to shrink by 50 percent. By the turn of the century, paper will satisfy less than 5 percent of the total commerce in information."[16] It is inevitable that, as technology becomes more sophisticated, products and services are replaced by new designs and processes. The printed book is a carrier of information just as the early 78 r.p.m. analog sound recordings were carriers of information. Throughout recorded history, the media that are used to carry information have changed as technology has developed more efficient, cost-effective means to store that information. The book is also destined to change in form: "The dollar relationship between

various methods of delivery for intellectual matter — be it information, education, or entertainment — must be recognized by the library profession. In terms of cold, hard cash, it is simply less expensive to distribute information electronically than by paper ONCE THE COMMUNITY IS EQUIPPED ELECTRONICALLY."[17]

The Sony Corporation recently introduced what might be considered the first "virtual book" in the form of the Bookman, a small, personal CD-ROM reader. Virtual books will take other forms as well. Raymond Kurzweil writes, "Virtual books will undoubtedly take many forms, but we can envision the basic model as a thin light slab with sizes ranging from pocket-sized to the full surface of one's desk. Resolution, color, contrast ratio, and lack of flicker will all match high-quality paper documents. These truly personal computers will be able to send and receive virtual books instantly through wireless communication."[18]

Books and libraries are in transition. The library profession must now face the inevitable — the library of the twenty-first century will be very different from the library of today. Those who complain that the traditional library is dying are correct. If those same people do not embrace the new technologies and take an active role in determining their applica-

tions in the library, they will surely be left behind. The development of electronic library services should not be left to the technocrats. Librarians must educate themselves in the use and application of new technologies and design information delivery systems in cooperation with computing and networking professionals. Consequently, graduate library programs need to revise their curricula to provide effective training in the the use of sophisticated technologies as well as the design of integrated information delivery systems.

The electronic library concept offers almost unlimited opportunities to devise new and innovative services. If we accept the challenge, the profession stands to gain immeasurably from the development of the virtual library. We have the opportunity to establish the direction of a major transition in librarianship and library service. The choice is ours.

References

1 H. L. Capron and Brian K. Williams, *Computers and Data Processing* (Menlo Park, CA: Benjamin/Cummings Publishing, 1982), 49-51.
2 Ibid., 52-53.
3 Ibid., 55-56.
4 Ibid., 57.
5 Tracy LaQuey with Jeanne C. Ryer, *The Internet Companion* (Reading, MA:

Addison-Wesley Publishing, 1993), 3-6.
6 Vannevar Bush, "As We May Think," *The Atlantic Monthly* 176 (July 1945): 102.
7 Ibid., 104.
8 Ibid., 106.
9 Ibid., 107.
10 Susan J. Barnes, "An Electronic Library Grows," *Computers in Libraries* 13: (September 1993): 12.
11 Ibid., 15.
12 Welcome to Project JANUS, the Columbia Law School Digital Library [Online]. (1993, November 30). Available FTP: **ftp.janus.columbia.edu** Directory: pub/general File: jan_info.asc.
13 Ibid.
14 Harriet Hagenbruch, "American Memory — History Meets the Age of Technology," *Library Software Review* 13 (Spring 1994): 35.
15 Michael Gorman. "The Treason of the Learned: the Real Agenda of Those Who Would Destroy Libraries and Books," *Library Journal* 119 (February 15, 1994): 130.
16 John Kountz, "Tomorrow's Libraries: More Than a Modular Telephone Jack, Less Than a Complete Revolution — Perspectives of a Provocateur," *Library Hi Tech* 40, 10 (1992): 39.
17 Ibid., 40.
18 Raymond Kurzweil, "The Virtual Library," *Library Journal* 118 (March 15, 1993): 54