

LJ INFOTECH

□ ONLINE DATABASES □

BY CAROL TENOPIR

Generations of Online Searching

[Based on a presentation to the Florida Library Association.]

OURS IS AN AGE of online searching, with new generations of systems and of users. With an estimated six million users of consumer online services such as America Online, CompuServe, and Prodigy and tens of millions of Internet users, online is now mainstream. Although the concept of retrieving information online is nothing new to librarians, there are new aspects: the look of the information, the amount of it available, the number of people it affects, and its impact on society.

However, even though the mainstream press implies that online is a product of the 1990s, most of what is happening today has its roots in earlier online generations. Today we are in the midst of what might be called the third generation of online. The first generation lasted until personal computers became ubiquitous—about 1981: searchers used dumb terminals; transmission speeds were slow; bibliographic databases ruled; intermediaries searched only commercial online systems such as DIALOG.

The second generation, from 1982 to 1991, began with PCs: they offered medium transmission speeds (up to 9600 bps), with bibliographic, directory, and ASCII full-text databases. Specialized end users joined intermediaries, but the general (consumer) end user was rare. Commercial online services coexisted with CD-ROM and locally loaded tapes.

Now, the third generation is characterized by multimedia PCs, much higher transmission speeds, and both ASCII and image full-text databases. As the Internet and consumer online services have grown, consumer end users

have joined intermediaries and specialized end users. The fourth generation is coming soon.

Evolution

From a social perspective and by sheer numbers, the impact of the Internet may seem revolutionary. These social changes are and will be far-reaching and dramatic; however, from a technological and an information resource viewpoint, the progression from first generation to third is evolutionary. Evolution does not mean that something new completely replaces or displaces the old. The best traits of the old are passed on to the new just as older species coexist with newer ones.

The three generations have evolved in five specific areas: 1) content/databases (WHAT); 2) people involved (WHO); 3) places people search (WHERE); 4) why they use online (WHY); and, 5) search engines/interfaces (HOW).

WHAT (content/databases)

The content of databases has evolved from bibliographic to ASCII full text to combination ASCII plus image to multimedia. Bibliographic databases are just finding aids, while ASCII full text can cheaply and quickly provide answers and complete texts. Image databases are facsimile replacements for paper documents, and multimedia go beyond the capabilities and uses of print. The next generation likely will add fact-answering databases, which provide specific answers, as well.

This evolution has had several impacts on libraries and online searchers. To begin with, hardware costs money. We need new and better computers, monitors, modems, printers, and CD-ROM drives. By the end of 1996 or early 1997 the new high-density CD-ROMs will be out. DVD (digital videodiscs) will require a new drive, although the new drives will also play old discs. Multimedia capabilities keep improving but require higher-end computers and faster modems to fully exploit them.

The web and consumer CD-ROM market are multimedia worlds. Electronic books and the new generation of

E-journals have color photos, sound, movement, and text. After using this all evening at home, when a patron comes to the library and gets plain ASCII text, it looks mighty old-fashioned. Most libraries are just beginning to tap the potential of online and CD-ROM multimedia services.

Some database services combine types to offer the best of both. For example, new online full-text services provide a bibliographic database for precise searching, then offer full-text document delivery online of requested texts or images. IAC SearchBank, UMI ProQuest Direct, and EBSCOHOST offer format choices with a different price per article depending on the form chosen.

Another continuing challenge is quality of content (not just appearance). We all have accessed beautiful homepages that are nothing more than junk. Web junk food may be appealing for a while but can lead to information malnutrition. Some may provide erroneous or biased information. The librarian's task of selection, evaluation, and pointing to authoritative and high-quality content is more important than ever.

Librarians must begin to document quality. If your boss asks, "Why are you paying \$10,000 a year to DIALOG or Lexis-Nexis or UMI when everything is there free on the Internet?" be prepared to show: 1) what is not on the Internet but is elsewhere; 2) what is on the Internet but is of better quality or is easier to use elsewhere; and 3) the good net sites you are using and the money you have saved by getting this information from the net.

WHO (people involved)

While searchers online were once intermediaries, that population has evolved to specialized end users (such as lawyers) and now general end users. All still coexist and will continue to do so. However, at each step in the evolutionary process the number of people searching increases dramatically, while their expertise often decreases. This impacts professional searchers in several ways. Users expect that more information can be found faster and cheaper but



Carol Tenopir is Professor at the School of Library and Information Science, University of Tennessee at Knoxville. Her E-mail address is tenopir@utkux.utk.edu

ONLINE DATABASES

may not have the skills to search most effectively.

Our role as teacher or instructor becomes vital as libraries offer an increasingly complex mix of commercial online systems, World Wide Web, CD-ROM, local area networks, and locally loaded databases. This has always been a role for librarians, but it predominates now in all information settings. Not only must we teach patrons how to access a variety of CD-ROM and online resources, we must also address larger ethical issues such as copyright, plagiarism, netiquette, and what to do with all the information retrieved.

WHERE (places people search)

Searching within one library has evolved to using the collections and resources of myriad library systems and searching from one's home, office, or dormitory. Even if users are located outside the library, they should be able to go through the library to access useful Internet sites or commercial online services such as FirstSearch.

Libraries and library services must go where the users are. Long before there is a virtual library there will be virtual library users. This will affect staff and staffing patterns. We need extended 24-hour help services that link reference librarians to online library users. Like the web, the online library becomes a 24-hour-a-day entity, but few libraries have taken the challenge of reference help at odd hours. Declining or steady state budgets may actually decrease service hours, rather than increase them.

WHY (uses for online services)

While people once searched online to create serious bibliographies from formal structured sources, that has evolved to self-service searching of many formal sources (with more emphasis on browsing), to highly interactive recreational use. Intermediary searchers used to get on and get off as quickly as possible; specialized end users began to browse more, and now general end users spend hours online.

The Find/SVP American Internet User Survey (published in January 1996) found that 9.5 million Americans use the web (3.6% of the population). Of the sites they visit, 52% are for personal reasons, 35% for business, and 13% for school. Only 11% use the web solely for business, and 29% never use it for business. The top application is E-mail, but

47% use it to retrieve product information, especially computer information. The average user spends 6.6 hours a week online.

The survey also asked for reactions to web use. Seventy-five percent said they have difficulty locating sites; a majority view it as a "selective reference source" rather than as an integral part of their daily media lives.

When respondents were asked to rank a series of activities that they believe will be prime movers in the future,

Users should be able to go through the library to access useful Internet sites or online services

Information Access came out on top with 70% and Communications next with 68%. Advertising/Marketing and Entertainment generated much less enthusiasm. [See: <<http://etrg.findsvp.com/features/newinet.html>>.]

These results have several implications for libraries. While online offers serious informational uses, the survey shows that online can supplement all the variety of a library's collection and purpose (recreational, educational, and social). However, queues may develop for a limited number of workstations as patrons use E-mail or play games. Public librarians are concerned that if the indecency provision of the 1996 Telecommunications Act is restored, they may be held responsible for what their patrons pull up online.

HOW (search engines/interfaces)

Search engines and interfaces have evolved from command interfaces with Boolean searching to menu interfaces with Boolean searching to graphical user interfaces (GUIs) or natural language interfaces with hypertext or relevance ranking. The next generation may include social agents (like Microsoft's Bob) or virtual reality (like the Holidock) natural language interfaces with linguistic search engines.

Interfaces have become more appealing and easier to use, but most still require some guidance and training. Interacting with multiple systems can be

confusing for anyone. Relevance search engines make searching more of a black box, with the system, not the searcher, doing most of the work. Hypertext and relevance ranking make it harder to pinpoint an exact search topic and often more difficult to tell why you are getting (or not getting) what you need.

Proliferation of software has improved recently as Netscape-type browsers have become a de facto standard and the common interface of choice. Many commercial online services are introducing web versions of their services this year.

Survival skills

How can information professionals survive as new generations of online evolve? First, maintain skills on all systems and all types of databases from every generation, because new resources supplement the old, rather than replace them. Library administrators must budget for training and allow staff time to learn the new techniques while keeping up with the old.

In the emerging Information Age, librarians must become excellent lease negotiators. Online full-text systems including desktop delivery systems such as Hoover and Lotus Newsstand are all priced in a lease/use arrangement. Librarians must learn to evaluate and negotiate these leases.

Also, we must work toward integrating all information media. Engineering Information Village (see "Moving to the Information Village," Online Databases, *LJ*, March 1, p. 29-30) is a good model for libraries to emulate. It guides users to useful web sites, provides online bibliographic databases, assists document delivery, and connects the user with librarians or engineers for help. It is a crucial service to have an information professional guiding users to the best sources regardless of the medium.

Finally, we must present results and instruction in terms that our customers will understand. Over 79% of large companies are now working toward intracorporate web sites—the web has become a familiar format, even to those who don't go online. Take advantage of this familiarity by presenting search results in that format, uploading directly to E-mail boxes, and rethinking user instruction with web metaphors in mind. The media hype about online may lack historical context, but it does provide a starting point for librarians to truly educate users.