10-1-1991

Predicting the Future

Carol Tenopir
University of Tennessee - Knoxville

ONLINE DATABASES

BY CAROL TENOPIR

Predicting the Future

[This column is based on a speech delivered at the 12th New Zealand Computer Conference, Dunedin, New Zealand, August 14, 1991]

TWENTY YEARS AGO, the first online databases became publicly available. Systems were small in 1972; DIALOG, for example, started with several subsections of the ERIC database. Then, as now, most database producers were government agencies, professional societies, or publishers that leased their information products to online services. Unlike now, online systems in the early 1970s were dominated by government-produced files, required direct dial access (there were no telecommunications networks), modem access speeds were slow, terminals were primitive, and users were few. The early databases were almost all bibliographic (indexing and abstracting).

In these 20 years we have seen many changes in the online database industry, most of which reflect steady growth and success. The number of online databases publicly available worldwide has steadily increased, from only a handful in the early 1970s to almost 5000 today. Although more than half of all databases are still produced in the United States, databases are now produced in over 34 countries and online systems in about 30. Not surprisingly, more people are conducting more searches and retrieving more information than at any time before.

The online industry has not been a complete success story. Many small database producers and online vendors receive very little use. Some major systems and databases dominate the online industry—few databases get used a great deal, many get used only a little. According to a 1986 article in *BusinessWeek* ("The Information Business," August 25, 1986, p. 82–90), "lots of companies find [the information business] alluring. But only a few are making money." They found that only eight companies—Reuters, Dun & Bradstreet, Quotron, TRW, Mead Data Central, Telerate, McGraw-Hill, and Dow Jones—made $100 million or more in electronic revenues in 1985. Still, the business continues to attract both large and small organizations.

Carol Tenopir is Associate Professor at the School of Library and Information Studies, University of Hawaii at Manoa, Honolulu

Online crystal gazing

The next few years are likely to bring even more dramatic growth and changes to all aspects of the online information industry. Some of these changes will result from technological factors; other are societal; still others, economic. Five major factors that will influence online database development in the near future are: 1) developments in telecommunications; 2) improvements in scanning and storage technologies; 3) expansion of database distribution options; 4) more complex needs and demands of a broader base of database users; and 5) changes in the dynamics of database production.

Developments in telecommunications—The first widespread telecommunications networks to provide inexpensive access to online systems became available in the mid-1970s. Today there are telecommunications network nodes in most cities in the world, linking users to almost all online systems. That's old news by now; what is changing is the speed and quality of communications via these networks.

In 1975 300 bps seemed fast compared to 110 TELEX access. Today most remote access online systems operate at 2400 bps (although many users still only have 1200 bps modems). To people accustomed to searching on local area networks (LANs), even 2400 bps seems incredibly slow. Recently a few online systems began offering 9600 bps access over voice-grade lines and more will follow. Telephone companies are switching to 56K bps with the ongoing conversion to digital phones. Not only will the access speed be much faster, but the transmission quality will be superior. Fiber optic transmission lines are being installed around the world, with major metropolitan areas generally first. Data rates and transmission quality will be going up dramatically in the next few years.

Higher access speeds with better quality will have some interesting impacts on the information industry and online users. Much online searching is done by reading and browsing texts once they are searched and retrieved. Human eyes cannot keep up even with 9600 bps, so the effective speed is much reduced as users pause to read and think. It may be more cost-effective to download massive amounts of information and postpone reading and thinking until after logoff—a practice that scares many producers. Once something is downloaded it can be transferred to a word processing or database program on the user's own computer. Conscience over possible copyright violations is all that prohibits a user from changing information or using it over and over again without paying further royalties.

Another impact has been changes in pricing algorithms. Although most online systems still charge by the connect hour, others have already changed to algorithms that base costs on characters transmitted, disk accesses required, or amount of information searched and retrieved. The new algorithms don't punish the user with a slower modem.

Higher access speeds will undoubtedly force all online systems to rethink their pricing policies. Differentials based on access speed, more reliance on charging for amount of information transmitted rather than connect hour, and flat fee contracts by the month will become more commonplace.

A negative impact of higher access speed has been unacceptable rates of line noise due to low-quality voice-grade telephone lines. Until fiber optic lines are in place, at 9600 bps, line noise may garble an entire

record. This problem is especially acute in countries with a poor telephone system infrastructure.

Higher transmission speeds allow more digital information to be transmitted at a reasonable cost. High-quality graphics can be transmitted in a reasonable amount of time, something that will dramatically change the appearance of online databases. Still mostly number- or text-based, online systems will incorporate more tables, charts, pictures, and other graphics. Equipment on the receiving end must, of course, be able to capture and print these bit-intensive transmissions.

Improved scanning and storage—Speedy transmission is not the only factor needed to enhance textual databases with graphics. Conversion of graphics in print products has not been easily done until recently. Better scanners and improvements in storage technology allow complete texts to be converted and stored. Coupled with higher rates of data transmission, online databases that more closely resemble print counterparts, but offer value far beyond print, will become common. Optical disc databases already contain pictures, motion, and sound, and online will feel the need to compete.

Improved storage technology at the online system end will mean not only more graphics online but larger and larger databases. Bibliographic databases with millions and tens of millions of records will become more common. Improved scanning coupled with massive amounts of storage and high-speed transmission will result in the rapid growth of full-text databases in the next few years. Full-text adjuncts to bibliographic databases will become the rule rather than the exception.

Increased storage capacity at the user's end will make downloading of massive amounts of information possible. So much information online or downloaded is not always positive. Searches that retrieve thousands of records or millions of full-text characters present the real problem of information overload. Database producers will need to face the issue of offering more quality filters to the data they present.

Expansion of distribution options—The online information industry has taken notice of the success of CD-ROM for database distribution. While it took many years to establish widespread use of online databases, CD-ROM has established itself in the library market in just a few years. Many academic libraries are reporting drastic reductions in the amount of online searching as they purchase more sources on CD-ROM.

On the whole, however, it appears that CD-ROM is providing database access to a new group of users; it is expanding markets rather than replacing them. Information Access Company reports, "Online revenues are certainly going up; there has been no impact in our traditional online market which is corporate libraries." According to Libby Trudell, marketing manager at Dialog Information Services, "in some cases, CD increases online use because it makes users more aware of electronic reference sources. We have seen some replacement, but there is not a consistent pattern."

One impact of CD-ROM has been the need for online systems to emphasize the relative benefits of online access over CD-ROM. The most up-to-date information is only available online and database producers and online systems are rushing to enhance update schedules. Wire services are updated continuously throughout the day and night on several online systems; other databases that used to be updated monthly are now updated weekly or even daily.

In the next few years, other optical media will open new distribution options to compete with online. Multimedia CD-ROM, CDTV, and CD-I all offer electronic publications that look much more attractive than ASCII text online. They offer high-resolution graphics, photographs, moving pictures, and sound in addition to text. They allow interaction with the contents and offer the user an opportunity to get involved with the learning process. These products are visually attractive and appeal to a much wider audience than online systems.

More users—More distribution options means more users of electronic information. As people get accustomed to different types of information, there will be a downturn in the reliance on printed sources and an increase in the users online. These new users will have different expectations and experiences with electronic information. Current interfaces will not be sufficient.

Better interfaces that accommodate a range of expertise will become more prevalent in the next few years. Online vendors already know that they cannot hope to attract a widespread end user market with the command-driven interfaces of the last 20 years. The option of menus is here now; graphic interfaces and more innovation and choices in interfaces will be coming online soon.

Development of expert online systems is ongoing and more will be available for access to the large publicly available online systems. Some of these expert system interfaces are sold by third parties as front-end communications software with value added. The major online systems will offer their own expert system interfaces as well in order to compete in the end user market. They will incorporate the search strategies and knowledge of experienced searchers as well as knowledge of the grammatical structure of the texts being searched.

Changes in industry dynamics—Online database use is a worldwide phenomenon, yet a majority of the large databases and online systems are produced by U.S. companies. But this is beginning to change; U.S. information companies are being purchased by non-U.S. corporations, even though production and offices may stay in the United States. Robert Maxwell now owns two of the popular online systems, ORBIT and BRS (Maxwell Online), in addition to Pergamon (a British firm).

The market is getting more global with the improvement in telecommunications networks and systems in countries around the world. With more reliance on satellite communications, online searching capabilities in many areas will improve.

Since many small database producers are not making money, market fallout will occur in the next few years. Some of the smaller companies will get out of the electronic information business altogether. The larger companies will get bigger as competition declines. The total number of databases, database producers, and online systems will continue to increase, however, as more traditional print information companies expand to electronic information. Much of this expansion will be in CD-ROM as well as online.