

ONLINE DATABASES

BY CAROL TENOPIR

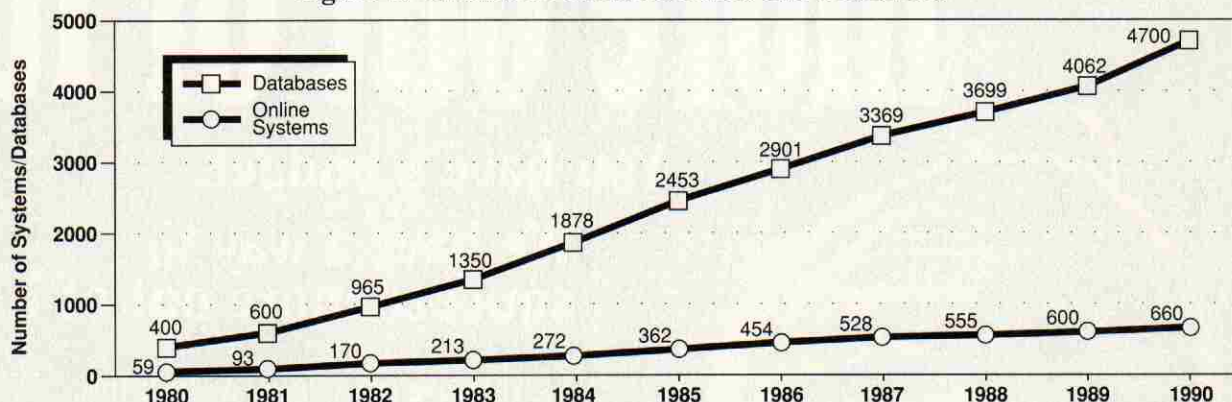
The Most Popular Databases

HOW MANY different databases do you search in a month? In a year? Do you now search many more and a wider variety of databases than you did ten years ago? At the rate online databases have proliferated in the last two decades, it is no wonder that we should all be searching a much greater variety than we used to. Where we once had a choice of three databases on a topic, we may now have a choice

the preface to *Computer-Readable Databases: A Directory and Data Sourcebook*, now published by Gale Research, and in the introduction to the *Proceedings of the National Online Meeting*, published by Learned Information. The most comprehensive source on database use is Williams's *Information Market Indicators (IMI)*, a quarterly publication aimed at database producers/vendors.

across several online systems, types of database producer, and subjects. IMI reports that corporate libraries and information brokers do the most searching. Overall figures, excluding law schools, law libraries, and legal firms (which are high-volume searchers), show the databases searched the most in 1990 were, in order, MEDLINE, NEXIS, ERIC, World Patents, Predicasts files, CA (Chemical

Figure 1: Growth in Online Services and Databases



Source: Cuadra, *Directory of Online Databases*, Cuadra/Elsevier, quarterly

of 30 or more. Yet a small number of databases still dominate the use in libraries and information centers.

Sources for information

The growth in the number of databases, companies that produce databases, and online services is tracked mainly by two long-time leaders in the field. Carlos Cuadra gathers much of his data for the quarterly *Directory of Online Databases* and semiannual *Directory of Portable Databases*, joint publishing ventures by Cuadra Associates and Elsevier. Martha E. Williams of the University of Illinois summarizes some of her data each year in

Growth of & most-searched databases

If you feel overwhelmed by the thought of all your choices in databases, a glance through these publications will show you why. In 1975 (when I started searching), Williams reports there were only 300 databases commercially available through about 100 vendors. By 1990, she found over 5600 computer-readable databases available from over 850 vendors. As seen in Figure 1, Cuadra counts similar growth numbers for online databases from 1980. For example, DIALOG has grown from just a handful of databases in 1975 to over 400 today.

With so many databases/online systems to choose from it is not surprising that the number of searches has increased dramatically. Williams reports a more than 30-fold increase in searches on word-oriented databases in the U.S. library/information center market, as shown in Figure 2.

The most popular databases cut

Abstracts) File, CA Search, and Financial Information Service.

High-volume corporate searchers skew these results somewhat. When academic libraries are considered alone, the picture changes a bit. The top databases in academic libraries in 1990, almost all bibliographic, were MEDLINE; ERIC; PsycInfo (online version of *Psychological Abstracts*); BIOSIS; Health Planning and Administration; CA File (Chemical Abstracts file with abstracts, available on STN International; Dissertation Abstracts; CA Search (Chemical Abstracts file without abstracts, available on several online systems); Inspec; CAB, Social Science Citation Index, PDQ—Physician Data Query, Catline, and ABI/INFORM in a virtual tie; National Newspaper Index; and Magazine Index. Public libraries do a relatively small percentage of the overall online searching, but tend to search these same databases. ERIC, MEDLINE, and Magazine Index get



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

CHARTS BY ANNABELLE CARTER

ONLINE DATABASES

high use in public libraries. A similar picture was found by Jamshid Beheshti and A. Tabah of McGill University in Montreal. They analyzed a random sample of libraries listed in the *Gale Online Databases Search Service Directory* to find which online systems and databases were most used by libraries and how these online services are being funded in the library. Fifty-one percent of their sample were academic libraries, followed by government libraries (21%), corporate libraries (21%) and public libraries (7%). Although the analysis was reported at the 1990 meeting of the American Society for Information Science (ASIS), the data were collected by Gale in 1987, so in most cases it represents pre-CD-ROM information.

Beheshti and Tabah found that

of these databases have print equivalents, CD-ROM versions, and/or microform versions. Librarians may use them in other forms and feel familiar with the content and idiosyncrasies of the product. They are almost all well-established bibliographic sources, popular in a multitude of formats in libraries for years.

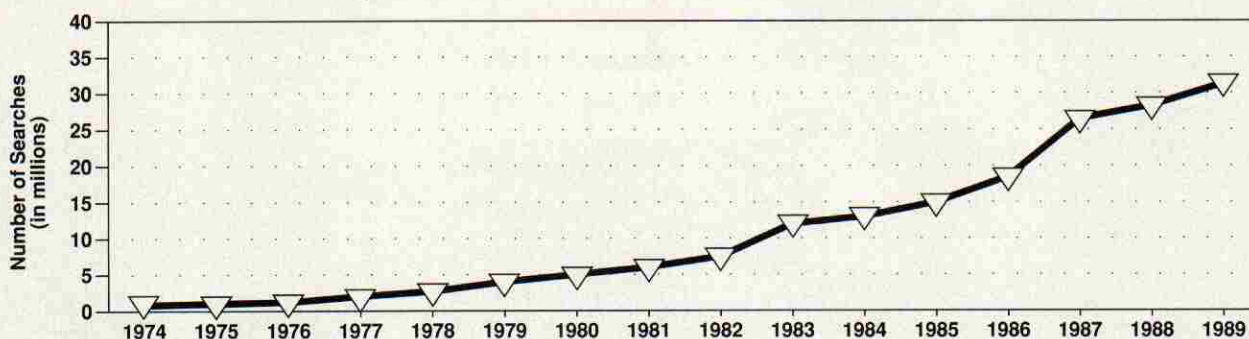
Another reason is *availability*. Most of the popular titles are available on multiple online hosts. Some, like MEDLINE, NTIS, and ABI/INFORM, are available on over ten online systems each. Most are available on several major online hosts. These databases carry the weight of *authority*. The database may be viewed as an authoritative or comprehensive source in a field, such as ERIC in education, CA Search in Chemistry,

bases that are on DIALOG, the average price per hour is approximately \$80. Still, affordability is clearly not a single deciding factor since INSPEC, Social SciSearch, and ABI/INFORM are all over \$100 per hour on DIALOG.

Popular CD-ROM databases

Academic libraries use their online search statistics to help make purchasing decisions for CD-ROM. It is not surprising that many of the same databases are popular on CD-ROM. Among OCLC libraries, the top ten CD-ROM databases as reported in the March 1990 issue of *Laserdisk Professional* are, in order, InfoTrac (some InfoTrac versions are CD equivalents to Magazine Index); ERIC; PsycLIT; Books in Print Plus; Academic Index; MLA Index; ABI/INFORM; DIS-

Figure 2: Growth in Online Services of Databases



Source: Williams, Computer - Readable Databases, Gale 1991

ten databases comprise 60% of all the databases used. Within the remaining 40% over 40 databases were mentioned. The ten most used databases in U.S. and Canadian libraries according to their calculations are ERIC; MEDLINE; PsycInfo; BIOSIS; ABI/INFORM; Compendex (Computerized Engineering Index); CA Search; Science Citation Index; and Health Planning and Administration.

Explaining database popularity

Why are some databases so popular? The first reason is *longevity*. Most of the databases that appear in several lists have been online for many years, e.g., ERIC, NTIS, CA Search, and MEDLINE were among the very first databases online. We are used to searching them, and their producers and vendors have years of experience to ensure their consistency and quality. The second reason is *familiarity*. Familiarity ties in with longevity, but goes beyond it. Most

MEDLINE in medicine, and BIOSIS in biology. Many offer one-stop shopping for a topic in a broad subject field, and some are megafiles with hundreds of thousands or millions of records.

Another common denominator is *structural quality*. Practically all of the most popular databases in libraries are bibliographic with controlled vocabulary descriptors. Many have other value-added fields such as classification codes. Thesauri and other database search aids are available for most. Professional searchers feel comfortable formulating and running searches when they have a variety of strategy options and can count on structural consistency. Many of these databases offer *affordability*: MEDLINE, ERIC, PsycInfo, and Health Planning and Administration are among the lowest-priced databases on major online systems. STN International offers substantial discounts for academic users using STN to access the CA files. For 13 of the 16 most popular data-

CLOSURE; Dissertation Abstracts; and Social Science Index.

In a July 1989 report, Williams summarized statistics on CD-ROM use. She found the most popular CD-ROM databases in a broader library and information center market to be ERIC, Books in Print Plus, MEDLINE, DISCLOSURE, InfoTrac, and PsycLIT. The top four CD-ROM vendors by expenditure are Information Access Company (InfoTrac), SilverPlatter, Lotus Corporation, and Wilson. Although fourth in revenues, Wilson is first in number of databases sold to libraries. It offers a large number of relatively low-priced CD-ROM databases.

The number of online and CD-ROM databases will continue to increase at an impressive rate. Most of our old favorites are bound to remain and will continue to be the backbone of our searching operations. Online and CD-ROM gives us the opportunity to meet new online faces as well.

Reed Business Information, a division of Reed Elsevier, Inc. All Rights Reserved.
Copyright of Library Journal is the property of Reed Business Information and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.