BIBLIOGRAPHIC DATABASE COMPARISONS

James W. Markham
Library - Cataloging Department
University of California, Santa Barbara
Santa Barbara, CA 93106-9010, USA

ABSTRACT

Bibliographic databases in online and CD-ROM versions and their corresponding abstract journals have been compared by numerous authors. A sample of these articles was analyzed to determine problems investigated and methodology used. Aspects compared can be placed into three categories: coverage, indexing, and database protocols/capabilities. Methods for comparing coverage and indexing involve searching databases and analyzing retrieved documents. Most search methods are either known item searches or subject profile searches. Advantages and disadvantages of these methods were compared. It is suggested that all three aspects should be compared for optimal evaluation of databases and that either search method could be used, but with precautions to avoid reported problems. Known item searches should search for recent articles of known value. Subject profile searches should use multiple search statements and employ descriptors as well as natural language.

INTRODUCTION

A large number of comparisons of bibliographic databases has appeared in the literature. Now that recent articles are comparing CD-ROM as well as online databases, the number of comparisons continues to increase. Nearly 10 years ago, Carol Tenopir (1982b) wrote, "Considering that the literature of database evaluation has been established for a decade, it is time to compare methodologies, test assumptions, and seek the best ways to evaluate database coverage." It has become evident that such a need exists more than ever, especially for science databases. In this study, a sample of database comparisons was surveyed in order to analyze their basic features and make some recommendations.

METHODS

Articles discussing database comparisons, including comparisons of print, online, and CD-ROM versions of various bibliographic indexes, were retrieved by various methods. Methods included an online search of the Dialog databases LISA

and Eric; a search of *Current Contents*, available on the University of California's MELVYL system; and scanning the last several years of *Online, Online Review, Database, the International Association of Aquatic and Marine Science Libraries and Information Centers (IAMLIC) Proceedings*, and the bibliographies of retrieved articles. Articles were rejected if the comparisons described were merely a repetition of the contents of user manuals for databases, without any tests, or if they did not provide details of methodologies used. A total of 42 articles remaining were analyzed to determine what problems were investigated and the methodology used.

RESULTS

Printed indexes are included here among the "databases" considered because some of the problems considered and methodologies used are the same as in comparing online or CD-ROM databases. Of the 42 studies considered here, 33 deal with online databases, three deal with CD-ROM databases, five deal with printed abstracts, and one compares online searching with manual searching in related printed indexes. All comparisons studied are cited in the bibliography at the end of this paper.

The aspects of databases compared in various studies appear numerous at first reading, but further analysis indicates that all of them can be placed into three categories, which are designated here as Coverage, Indexing, and Database Protocols/Capabilities.

Coverage

Most of the studies considered here used coverage as a basis of comparison of databases. At least seven other factors were tested which are considered here to be aspects of coverage. These are overlap, uniqueness, content, relevance, novelty, journal coverage, and currency.

Overlap is important as a measure of how much databases have in common, and 19 of the studies specifically mentioned overlap (e.g. Bearman & Kunberger, 1977; Datta, 1988; Yonker, *et al*, 1990). Gluck (1990) reviewed overlap in detail and proposed a new definition involving additional statistical measurements.

The opposite of overlap is uniqueness, a measure of how much of the content of a database is not shared with any other (Meyer *et al*, 1983; Snow & Ifshin, 1984; Brandsma *et al*, 1990).

Content measures the emphasis of the database (Hightower & Schwarzwalder, 1991) or what kind of literature is in the database.

Relevance of retrieval is a measure of the value of a database and is used in the calculation of precision, or percentage of relevant documents in the retrieval. It involves judgement of the retrieval by the searcher or another subject expert (e.g.

Markham, 1990). In some cases relevance has been called "utility" (Snow & Ifshin, 1984), which measures additionally whether the relevant articles can be used, e.g. are they in English, etc.

Related to relevance is novelty, a measure of how many retrieved documents are new to the requestor. This may not reflect the value of the database very accurately but rather may reflect only the requestor's familiarity with the literature (McCain, *et al*, 1987).

Journal coverage is a special aspect of coverage. It provides a useful means of comparing databases. Some earlier studies, not included here, compared journal coverage after merely polling database producers to obtain a list of journals abstracted. Journal coverage does not measure the same thing as coverage in general, because a journal may be listed as "covered" even if only a few articles from a journal were selectively indexed by a database. (See also discussion by Gluck, 1990)

Currency, also called timeliness, is a measure of the time lag between publication of a document and referencing of the document in a database (Brooks, 1980).

Indexing

The relative quality of indexing has been used by many authors to compare databases. It is a valuable measure for comparison, but there is considerable disagreement on how it should be measured (e.g. Sievert, 1987; Chu & Ajiferuke, 1989).

Database Protocols/Capabilities

Many reports comparing search protocols or database capabilities are merely listings of search commands or reiteration of the contents of search manuals and were not considered here. Of those considered here, Brooks (1980) considered online search capabilities as well as coverage of 4 online databases; Hewison (1989) included "search functions", "output capabilities" and "user-friendliness" among the items to check in evaluating CD-ROM versions of Medline. Preuss (1990) compared retrieval by searching the same request online and in related printed indexes. This compared search capabilities of the two methods. It also turned up a notable lack of overlap in retrieval by the two methods.

Methodology

Methods for comparing coverage and indexing involve searching the databases and analyzing the resulting retrieved documents. The search methods fall into two basic categories, "known item" searches and "subject profile" searches. Each of these

has several versions.

Known item searches have been termed the "bibliography method" by Tenopir (1982b). The lists of known items to be searched for have been derived from several sources. The commonest source is a bibliography or review, which are not the same. "Citations that have been systematically chosen by a reviewer in the field imply some quality judgement ... ; citations chosen by a bibliographer may not imply quality judgement", (Tenopir, 1982b). Other sources of known items include publications lists of subject specialists (Montgomery, 1973; Sokolov, 1988); papers selected randomly from core journals for a discipline (Haas, 1991); all papers from randomly selected journals (Jérôme, 1973; Starr, 1982); all publications from a certain date (Jérôme, 1973; Cloutier, 1991); all retrieval of articles from certain journals in a large database which was then searched in a smaller database (Sievert & Verbeck, 1987); or lists of highly co-cited documents (White & Griffith, 1987). Known items can be searched by any method to determine their presence or absence; the commonest approach is an author search for articles.

Subject profile searches have several versions, each of which has been used by several authors. The basic approach is to use the same free text search (also called natural language) in all databases (e.g. Tillotson, 1988; Bac, 1990; Hiller, 1991). The free text search may or may not involve terms which are used as descriptors in certain databases. An identical descriptor search can be made across databases (e.g. McCain, *et al*, 1987; Datta, 1988; Stern, 1990). Several different search statements can be run across several databases to compare (e.g. Meyer, *et al*, 1983; Young and Minion, 1988; Markham, 1990). The same combined search containing descriptors and free text can be run in several databases (e.g. Soremark, 1990; Stern, 1990). Finally, the same topic can be searched across databases using searches tailored to each database (Watkins, 1981; Corbett & Ifshin, 1983).

Currency has been measured by retrieving a set of items input in a certain year, then intersecting this set with sets of records published in that year and earlier years (Hightower & Schwarzwalder, 1991). It can also be measured by searching for newly received items at various time intervals.

Indexing has been compared in various ways. The most common is to examine the indexing assigned to retrieved documents (e.g. Bac, 1980; Starr, 1982; Sievert & Verbeck, 1987; Soremark, 1990) or to compare this with the indexing of known items not retrieved (McCain, *et al*, 1987). The examination may result in subjective judgement about indexing quality. White and Griffith (1987) detailed methods for calculating rarity and discriminant value of descriptors. Alternative calculations were detailed by Chu and Ajferuke (1989). Indexing has also been tested by comparing retrieval using various descriptors or types of descriptors or by comparing retrieval using descriptors vs. free text (Watkins, 1981; Anholt & Hurt, 1988; Kooijman-Tibbles, 1989; Markham, 1990). Numerous studies have pointed out a need for additional index terms or categories of index terms in one or several

databases after testing retrieval (e.g. Bac, 1980; Brooks, 1980; Markham, 1990).

DISCUSSION

The aspects of databases which are used for comparison have been placed into three categories: coverage, indexing, and database protocols. All of these are valid measures for comparing databases and for optimal estimation of database quality, probably all three should be examined. Some databases with excellent coverage, especially some CD-ROM databases, may lack optimal search capabilities or be too difficult to use, or may have inadequate or inconsistent indexing. They would receive very different quality reviews depending on what was tested.

Tests of coverage can be carried out by known item searches or by subject profile searches. Tenopir (1982b) compared the two methods in detail, listing advantages and disadvantages of each, and recommended the subject profile approach. She noted that the known item or bibliography method may not give an up-to-date comparison because it is based on bibliographies which give a picture of the past, due to the time required to prepare bibliographies. The bibliographies in reviews may reflect the bias of the reviewer. This method is more costly and takes more time, since large numbers of references must be searched for. The subject profile method is much faster and less expensive and is not biased in the same ways. However, this method can be influenced by indexing policies of the database and it can be difficult to achieve optimal recall and precision rates so that valid conclusions about coverage can be drawn.

More recent studies have found ways to eliminate some of the disadvantages of both methods. Haas (1991) searched for recent articles from core journals for a discipline, thus eliminating the problem of the out-of-date bibliography as well as the bias of the review article compiler. Markham (1990) avoided some of the problems of indexing by running several different search statements using various combinations of controlled vocabulary and natural language. Hightower and Schwarzwalder (1991) avoided indexing problems by searching only in the title and abstract fields.

It appears that there is still disagreement over methods to measure the value of indexing and no recommendation is made here.

The known item method requires more time, but gives more detailed information on coverage. The subject profile method is faster but is harder to design to give a valid picture of coverage. It appears that neither method should be favored over the other, but rather that either, or perhaps both, methods should be used, but with precautions to avoid the noted problems. The known item method should search for recent articles of known value for the discipline of interest. The subject profile method should use more than one search statement and should employ descriptors as well as natural language in various combinations. This method also allows a test of indexing by measuring how the use of probable index terms improves recall.

BIBLIOGRAPHY

Anholt & Hurt, H. 1988. Undergraduate access to bibliographic databases: a comparison of descriptor dependent batch searching and online searching of ERIC. *Canadian Journal of Information Science* 13(3/4):34-39.

Bac, R. 1980. A comparative study by the PDR of toxicology information retrieval from online literature databases." *Online* 4(2):29-33.

Bearman, T.C. and **Kunberger, W.A.** 1977. *A study of coverage overlap among fourteen major science and technology abstracting and indexing services*. National Federation of Abstracting and Indexing Services : Philadelphia, PA.

Brandsma, R., Deurenberg-Vos, H.W.J., Bakker, S., Brand-de-Heer, D.L., Otten, R.H.J. and **Pinatsis, A.** 1990. A comparison of the coverage of clinical medicine provided by BIOSIS Previews and Medline. *Online Review* 14(6):367-375.

Brooks, K. 1980. A comparison of the coverage of agricultural and forestry literature on Agricola, BIOSIS, CAB and SciSearch. *Database* 3:[March] 38-40.

Chu, C.M. and **Ajiferuke, I.** 1989. Quality of indexing in library and information science databases. *Online Review* 13(1):11-35

Cloutier, K.A. 1991. A comparison of three online Markush databases. *Journal of Chemical Information and Computer Science* 31:40-44.

Corbett, P.K. and **Ifshin, S.L.** 1983. Online retrieval of environmental and occupational health literature: a comparative study. *Medical Reference Services Quarterly* 2(3):25-36.

Datta, V.K. 1988. Coverage of literature on mycotoxins by computer databases and comparison of this coverage with the TDRI in-house index facility. *Quarterly Bulletin of the International Association of Agricultural Librarians and Documentalists* 33(2):61-78.

Gluck, M. 1990. A review of journal coverage overlap with an extension of the definition of overlap. *Journal of the American Society for Information Science* 14(1):43-60.

Haas, S. 1991. Gaps and overlaps: A comparison of three water related CD-ROMs. IN: *Breaking the Barriers to the Free Flow of Information: Proceedings of the 16th Annual Conference, 1990.* Editors: Elizabeth Fuseler McDowell and Stephen Wiist. [s.l] : International Association of Marine Science Libraries and Information Centers. Fort Collins, CO. pg. 129-36.

Hewison, N.S. 1989. Evaluating CD-ROM versions of the Medline database: A checklist. *Bulletin of the Medical Library Association* 77(4):332-336.

Hightower, C. and **Schwarzwalder, R.** 1991. A comprehensive look at materials science databases. *Database* 14(2):42-53.

Hiller, S. 1991. CD-ROM sources for marine geology physical oceanography bibliographic information. *Geoscience Information Society Newsletter* 128:10-11.

Jérôme, S. 1973. A comparative study of the coverage of physics journals by two computerized data bases -- SPIN (Searchable and Physics Information Notes) andCAC (Chemical Abstracts Condensates). *Information Storage and Retrieval* 9:449-455.

Kooijman-Tibbles, H.C. 1989. Database evaluation: A case study from the Netherlands. *International Journal of Information Management* 9:179-185.

Markham, J.W. 1990. Online retrieval strategies and database comparisons for literature on macroalgae. IN: *LAMSLIC at the Crossroads: Proceedings of the 15th Annual Conference, 1989.* Editors: Robert W. Burkhart and Joyce C. Burkhart. [s.l.] : International Association of Marine Science Libraries and Information Centers : St. Petersburg, FL. pg. 157-61.

Martyn, J. 1967. Tests on abstracts journals: Coverage overlap and indexing. *Journal of Documentation* 23(1):45-70.

McCain, K.W., White, H.D., and **Griffith, B.C.** 1987. Comparing retrieval performance in online databases. *Information Processing and Management* 23:539-553.

Meyer, D.E., Mehlman, D.W., Reeves, E.S., Origoni, R.B., Evans, D., and **Sellers, D.W.** 1983. Comparison study of overlap among scientific databases in searching pesticide information. *Online Review* 7(1):33-43.

Montgomery, R.R. 1973. An indexing coverage study of toxicological literature. *Journal of Chemical Documentation* 13(1):41-44.

Preuss, B. 1990. Konventionelle Literatursuche versus Online: Literaturrecherche in den Datenbanken PSYNDEX und PsycINFO. *Nachrichten für Dokumentation* 41(6):361-368.

Price, C. and **Burley, R.** 1986. An evaluation of information sources for retrospective literature searching on occupational diseases. *Journal of Information Science* 12(5):257-265.

Sayers, M., Joice, J. and **Bawden, D.** 1990. Retrieval of biomedical reviews: A

comparative evaluation of online databases for reviews of drug therapy. *Journal of Information Science* 16:321-325.

Sievert, M.E. and **Verbeck, A.** 1987. The indexing of the literature of online searching: A comparison of ERIC and LISA. *Online Review* 11(2):95-104.

Snow, B. and **Ifshin, S.L.** 1984. Online database coverage of forensic medicine. *Online* 8 [March]:37-43.

Sodha, R.J., Kostrewski, B.J., and **Schier, O.** 1990. JICST-E (Japan Information Centre of Science and Technology--English File): its construction in providing access to online Japanese biomedical information with selected comparisons to EMBASE and MEDLINE. *Journal of Information Science* 16:93-98.

Sokolov, B. 1988. Arctic/Alaska gray literature: evaluation of access via commerical databases. *Glaciological Data* 22:263-80.

Soremark, G. 1990. MEDLINE versus EMBASE: Comparing search quality. *Database* 13(6):66-67.

Starr, S. 1982. Databases in the marine sciences. *Online Review* 6(2):109-125.

Stern, D. 1990. Artificial intelligence databases: A survey and comparison." *Database* 13(4):19-24.

Tenopir, C. 1982a. Distribution of citations in databases in a multidisciplinary field. *Online Review* 6(5):399-419.

Tenopir, C. 1982b. Evaluation of database coverage: A comparison of two methodologies. *Online Review* 6(5):423-441.

Tillotson, J. 1988. Databases for arctic marine technology. *Glaciological Data* 22:101-8.

Walker, 1990. Searching the humanities: Subject overlap and search vocabulary. *Database* 13(5):37-46.

Watkins, S.G. 1981. The IRL Life Sciences Collection and BIOSIS: A comparison of on-line access to the literature of biology. *Database* 4 [Sept]: 39-46.

White, H.D. and **Griffith, B.C.** 1987. Quality of indexing in online data bases. *Information Processing & Management* 23(2):211-224.

Wilkinson, D. and **Hollander, S.** 1973. A comparison of drug literature coverage by Index Medicus and Drug Literature Index. *Bulletin of the Medical Library Association* 61(4):431-432.

Wood, F., Berrie, A.T., Plampin, H.R. and **Wilkinson-Tough, M.L.** 1989. Evaluations using test queries of chemical hazards databases and databanks. *Journal of Information Science* 15:269-276.

Wood, J.L., Flanagan, C. and **Kennedy, H.E.** 1973. Overlap among the journal articles selected for coverage by BIOSIS, CAS, and Ei. *Journal of the American Society for Information Science* 24:25-28.

Yonker, V.A., Young, K.P., Horwitz, S. and **Cousin, K.** 1990. Coverage and overlaps in bibliographic databases relevant to forensic medicine: a comparative analysis of MEDLINE. *Bulletin of the Medical Library Association* 78(1):49-56.

Young, M. and **Minion, R.** 1988. A study of overlap between geographically oriented databases and subject oriented databases. *Glaciological Data* 22:227-34.