DOCUMENT RESUME

ED 368 362 IR 054 896

AUTHOR Wells, Amy Tracy
TITLE A Comparative Analysis of ERIC and LISA with an Emphasis on Database Composition and Authority Control as They Relate to Retrieval.
PUB DATE 17 Jul 93
NOTE 13p.
PUB TYPE Viewpoints (Opinion/Position Papers, Essays, etc.) (120) -- Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Authority Control (Information); *Bibliographic Databases; Clearinghouses; Comparative Analysis; Educational Research; Information Processing; *Information Retrieval; Library Education; Library Research; *Library Science; Relevance (Information Retrieval); Search Strategies; User Needs (Information)
IDENTIFIERS Database Development; Database Overlap; *ERIC; *Library and Information Science Abstracts

ABSTRACT

The ERIC and LISA (Library and Information Science Abstracts) databases are compared through a literature review and the analysis of an online search on a topic in librarianship. ERIC emphasizes education and includes serials, monographs, collections, and research reports for their coverage of some aspect of education. LISA is dedicated to topics in the field of library science and such related areas as publishing and bookselling. In the final analysis, ERIC was a superior source for information even though its scope is broader than library issues. A comparative analysis of coverage that included acquisition policy, types of material, currency, duplication, recall and precision, price, and the use of authority controls (thesaurus terms versus subject headings, specificity, exhaustivity, searchable field, and special problems) rated ERIC as better than LISA overall for library information. For some very specific topics, however, LISA might be the better source. (Contains 6 references.) (SLD)

A COMPARATIVE ANALYSIS OF ERIC AND LISA WITH AN EMPHASIS ON DATABASE COMPOSITION AND AUTHORITY CONTROL AS THEY RELATE TO RETRIEVAL

Amy Tracy Wells
17 July 1993

A COMPARATIVE ANALYSIS OF ERIC AND LISA WITH AN EMPHASIS ON DATABASE COMPOSITION AND AUTHORITY CONTROL AS THEY RELATE TO RETRIEVAL

This study is limited to a comparative literature review and analysis of an online search in DIALOG on a topic in librarianship. It is divided into two main parts, the first examining ERIC's (Educational Research Information Center) and LISA's (Library and Information Science) database composition and coverage with the second considering the issue of authority control and presenting the conclusion.

I. DATABASE COVERAGE

Users, prior to beginning an online search must first form search statements and consider the logical relevant concepts. At this point database selection can begin (or, of course, a search in DIALINDEX or OneSearch can be executed). As there is such variety in the individual databases, understanding their composition is necessary.

Ernest et. al. define ERIC as,

> ... two components: CIJE, which encompasses 750 journals (including at least 26 library periodicals), with 18,000 to 20,000 citations announced each year, and RIE, which indexes conference proceedings or unpublished papers often having limited or no other distribution. RIE adds 13,000 to 15,000 citations annually. [1]

Through a series of 16 Clearinghouses, specializing in various topics in education, ERIC seeks to capture and disseminate copies of all significant documents. It maintains a seemingly simple acquisition policy, by agreement it receives approximately half of its publications through organizations such as the Department of

Education or American Library Association. The rest, from conference proceedings to ephemeral materials, it receives through informal arrangements. This remaining material has an "...average acceptance rate of 50%." [2]

ERIC's coverage policy is easy to understand, as the emphasis is on education, the inclusion of all serials, monographs, collections, research reports etc. is dependent on their involving some aspect of education. This, naturally, includes their choices in Library and Information Science. It also emphasizes domestic coverage with foreign materials averaging only 2-3%. [3]

LISA, as its name suggest, is dedicated to topics within "...the field of Library and Information Science as well as such related areas as publishing and bookselling." [4] Like ERIC its materials include serials, monographs, research reports, theses, dissertations etc. In contrast, LISA indexes some 300 [5] journals and its annual citations average some 1,300. [6] Organizationally it is less complex than ERIC and also unlike ERIC it is decidedly international in its perspective. Published in Britain it is also known as the (British) Library Index. Ernest et. al. established a coverage of 44% of its materials to be non-United States, Canada or British publications with a 29% concentration of non-English titles. [7]

Even given this "policy selection" it is incorrect to assume that should a particular publication be indexed that it would be indexed in its entirety. That is to say, a feature piece from a given source, may appear but no other part of the source be

represented. Having noted this obstacle to comprehensive indexing there is further diversity or as Stieg notes,

> Within these boundaries selection can be described as capricious. Some interviews are indexed, some aren't; some short articles are, some aren't. [8]

Involved in the issue of selection policy is the issue of currency. While there are numerous reasons for conducting an online search, probably the main one is to seek the latest possible information. As a standard practice ERIC is updated monthly but as Ted Brandhorst, Director of ERIC Processing and Reference Facility, states, "...the best a document could do from acquisition to announcement would be 2 months. Average time is more like 3-4 months." [9] Complicating this time analysis is ERIC's practice of informal agreements for complimentary publications. 50% of its collection is under no formal time constraints, a decided disadvantage. These materials could be coming from a domestic source or from an international source and in foreign languages, another impediment. Yet considering all this ERIC's actual time is closer to 7.6 months. [10]

While LISA's selection policy is more standardized, its international coverage impedes its currency. Since 44% of its collection are foreign language/foreign materials they must also wait on the international mail system in addition to experiencing a time lag for translation purposes. In comparison with ERIC, LISA's currency is quite low, 10.3 months. [11] While one can appreciate the troubles associated with foreign mail service and

translating, still if currency is important this analysis does not recommend LISA.

The literature contradicts itself with regard to duplication of the two databases. LaBorie and Steig come to the conclusion that considerable overlap exists [12] and [13] while Ernest et. al. find the lack of duplication "startling". [14] Other areas of interest are recall/precision and pricing. When selecting a database these two issues become quite important. Beginning in a more expensive database and conducting an inefficient search has obvious cost drawbacks. Here LaBorie's finding was intriguing. An initial search conducted first in ERIC and then in LISA for relevant citations revealed a 70-30% split respectively. A reversed search revealed a 60-40% split. [15] The price differences amongst the databases is about half, in ERIC's favor.

II. AUTHORITY CONTROL

The second area to consider before conducting an online search is that of authority control. Would the search be more successful if controlled vs. natural language is used? How specific and exhaustive are the terms used? To what extent can certain fields be searched? How can a comprehensive search be executed? These are all questions which the online searcher may need to consider.

Here is another area of divergence amongst the databases. ERIC uses a hierarchical thesaurus while LISA utilizes subject headings. Each proved to have advantages and disadvantages. The online search

was seeking literature analyzing user error in an academic setting with CD-ROMs. In this case, neither presented see or see also references between "college library", "academic library" or "university library". The greatest inconsistency was in LISA's use of "users (information)", "user (needs)" and "user satisfaction (information)" and etc. to cover the same concept. ERIC also allowed searching by identifier, or use of natural language in conjunction to controlled language. (Natural language can be used in both.)

A significant difference was found in how the two databases addressed the areas of linking and chain linking, specificity and exhaustiveness. ERIC's uses of linking included items such as "tables (data)" to assist specificity whereas LISA uses chain linking to confusion, such as "computerized information storage and retrieval". A quantitative and qualitative analysis of ERIC and LISA's indexing by Sievert and Verbeck found LISA to be more exhaustive. They also considered specificity. Their definition of "unique" excluded 1) "grammatical or syntactical variation", 2) narrower versions of the same concept (ie. "library services" equated "reference services"), 3) synonyms and 4) "population or discipline" (ie. "higher education", "college library", and "undergraduates" were "...considered a single concept..."). [16]

While their definitions of what constituted "unique" items followed logically, their examples for item four did not. A reproduction of their analysis (utilizing data from the online search: nine retrieved items from ERIC and six retrieved items from

LISA) while redefining "population or discipline" yielded the following results:

	ERIC	LISA
Number of descriptors per item	10.80	14.57
Number of words (excluding "of" and "and") per item	21.80	30.42
Number of "unique" terms	9.66	8.86

Although LISA seemed to have greater specificity and exhaustivity the opposite is in fact true. LISA even included exactly identical descriptors in the same item 5 times.

In terms of searchable fields, an examination of the Blue Sheets yielded more similarity than diversity. The differences are therefore summarized below

ERIC		LISA
ID AV CN CP DT GL JA PN RN SP TA ZZ	similar to " "	FS RT, QL, FS
		ED PD RF SF SO ST

Both used the same name search strategy. ERIC allowed Limiting by /ED, /EJ, /MAJ, /MIN and LISA by /CR, /NONCR, /ENG, and /NONENG. Their formats were also quite similar with only two small differences noted between Formats four and six. [17] and [18]

Stieg and Atkinson provide examples in errors of consistentency, ie. William A. Katz being alternatively indexed as" B.","Bill", and "Bill (ed.)", alternate indexing of " Stieg" as "Steig", and the use of the British "Library Stock" unknown to Americans who use "Library Collection". They also note that of the two, ERIC made less of these types of errors. [19]

In the final analysis ERIC proved itself to be a superior source for information even in the field of Library and Information although LISA is the largest. A comparative analysis of coverage including acquisition policy, types of material, currency, duplication, recall/precision, and price and the use of authority controls including thesaurus terms vs. subject headings, specificity, exhaustivity, searchable fields, and special problems combined to rate ERIC, overall, as above LISA. Again, it is important to note that this conclusion is based, in part, on an analysis of a highly specific search.

When this analysis was begun, the anticipated database of choice was LISA. Due to the above considerations in a future search with the same type of search needs, ERIC would be the database of choice. However, in a search seeking information on public librarianship ERIC might not yield the best results. [19] Also important to consider in choosing a database might be the specific limiting capabilities which would otherwise render a search useless. For example, LISA's FV index might allow search to trace an important development not readily searchable through ERIC. In closing, ERIC though not specific to issues in Librarianship met

with a highly positive analysis.

TERM GROUPING

(PHYSICAL LOCATION) academic library, college library, university library
blind and partially sighted
comparative analysis, concept analysis
(INTERACTIVE SYSTEMS) computer networks, online systems, management information systems, online systems, integrated library systems, library automation, online catalogs, local area networks
cdrom, cd-rom, optical data disks, optical disks, discs, compact discs
computer software, computer software development, programs (computers), screen reading software
computer assisted instruction, programmed tutoring
cost effectiveness, costs
critical thinking, thinking skills
(STATIC OR INTERACTIVE) databases, data bases
error patterns, problems
document delivery, availability documents
financial support
foreign countries
france
futures (of society)
higher education
information needs, information seeking
information scientists
information services
information storage and retrieval, information retrieval, computerized information storage and retrieval, data collection
information technology
intermode differences
japan
library collections
library personnel
library role
library schools
(LIBRARY ORIENTED) library services, reference services, library instruction
library skills
medical libraries
medicine
(SINGULAR) microcomputers
online searching
output
policy formation
reaction time
reader services
research

search strategies
serials, periodicals
services
speech synthesizers
statistics
subject indexing, computerized subject indexing
tables (data)
telecommunications, data transmission
training methods
use
user needs (information), users, user satisfaction (information), users (information), user satisfaction
videotape recordings, storage media, electronic media

1. Ernest, Douglas J., Holley R. Lange, and Della Herring "An Online Comparison of Three Library Science Databases," RQ, 28, 2, Winter 1988, pp. 185-194.

2. Lopez, Manuel D., "ERIC Sphinx or Golden Griffin?", Urban Education, 35, 3, October 1989, pp. 287-306.

3. IBID.

4. Dialog Information Retrieval Services, LISA Blue Sheets, July 1988, pp. 61-1 through 61-4.

5. LaBorie, Tim and Michael Halperin "The ERIC and LISA Databases: How the Sources of Library Science Literature Compare," Database, 4, 3, Sept. 1981, pp. 32-37.

6. Dialog Information Retrieval Services, LISA Blue Sheet, Sept. 1988, pp. 61-1 through 61-4.

7. Ernest, Douglas J., Holley R. Lange, and Della Herring "An Online Comparison of Three Library Science Databases," RQ, 28, 2, Winter 1988, pp. 185-194.

8. Stieg, Margaret F. & Joan L. Atkinson "Librarianship Online: Old Problems, No New Solutions," Library Journal 113, 16, pp. 48-59.

9. Brandhorst, T. and J. Eustace [eds.] (1980) Indexing ERIC Processing Manual, Section 7. Washington, D.C.: Educational Resources Information Center. (ERIC Document Reproduction Service no. ED 219-088).

10. Ernest, Douglas J., Holley R. Lange, and Della Herring "An Online Comparison of Three Library Science Databases," RQ, 28, 2, Winter 1988, pp. 185-194.

11. IBID.

12. LaBorie, Tim and Michael Halperin "The ERIC and LISA Databases: How the Sources of Library Science Literature Compare," Database, 4, 3, Sept. 1981, pp. 32-37.

13. Stieg, Margaret F. & Joan L. Atkinson "Librarianship Online: Old Problems, No New Solutions," Library Journal, 113, 16, pp. 48-59.

14. Ernest, Douglas J., Holley R. Lange, and Della Herring "An Online Comparison of Three Library Science Databases," RQ, 28, 2, Winter 1988, pp. 185-194.

15. LaBorie, Tim and Michael Halperin "The ERIC and LISA Database: How The Sources of Library Science Literature Compare," Database, 4, 3, Sept. 1981, pp. 32-37.

16. Sievert, Mary Ellen and Allison Verbeck, "The Indexing of the Online Searching: a Comparison of ERIC and LISA," Online Review, 11, 2, April 1987, pp. 95-104.

17. Dialog Information Retrieval Services, ERIC Blue Sheet, February 1987, pp. 1-1 through 1-4.

18. Dialog Information Retrieval Services, LISA Blue Sheet, Sept. 1988, pp. 61-1 through 61-4.

19. Stieg, Margaret F. & Joan L. Atkinson "Librarianship Online: Old Problems, No New Solutions," Library Journal, 113, 16, pp.