

SHOPPING AROUND FOR MICROCOMPUTER SOFTWARE IN ONLINE DIRECTORIES — OPTIONS VERSUS WISH LIST OF A SEARCHER

P. Jacsó, SzAMALK, Hungary

Keywords: *Software Directories, Database Comparison.*

Abstract: Online software directories are to provide exact, factual data about computer programs and to make them accessible by the combination of several topical and technical criteria. The paper reviews how these requirements are met by the International Software Database, the Business Software Database and the Online Microcomputer Software Directory and compares them by size, currency, completeness, searchability and record structure.

1 INTRODUCTION

Our computer science library and information center has been providing information about computer software for a long time from printed directories and catalogs. These sources were far from ideal because of their inherent limitations. Their quarterly updates, yearly addenda were outdated in the moment they were received. Furthermore, a complete search required the inconvenient consultation of several back-issues, and the directories, catalogs were accessible only by the most common search criteria, like broad subject category, type of hardware and operating system required, name of software. Availability of some directories and catalogs in online version were to offer considerable improvement in searching for software, providing:

- up-to-date information on new software and new releases of existing computer programs
- one comprehensive source
- several added access points
- easy searching with combination of multiple criteria

Online abstracting/indexing databases with computer science coverage, like: Microcomputer Index, INSPEC, the Computer Database, Microsearch provide information about software very adequately when the search criteria relate to type of application and functional capabilities. Directory type databases are expected to provide searchability by additional, mostly technical criteria, like operating system and memory requirements disk size, price, backup facility, kind of support provided, etc.

This author has used three online software directories under two online systems:

International Software Database (MENU)	DIALOG
Business Software Database (BSD)	DIALOG, BBS
Online Microcomputer Software Directory (SOFT)	BRS

and gathered some experience which is reported in this paper to help fellow researchers and to plead with database producers and online system operators

for adjusting the structure and content of the records and of the index file to meet more adequately the requirements of software searching.

The three databases are compared in terms of subject coverage, size, currency, searchability and record structure. Sample records describing the Wordperfect software family are reproduced in Figures 1-7 from all the three databases to illustrate the comments made in the paper.

2 SOURCES COVERED, SIZE AND CURRENCY

The three databases provide coverage for different kind of software in terms of application and of computer category. This is summarized in Table 1.

1. Data referring to the size are valid as of July, 1986.

Database	Records	Computer classes and application categories
SOFT	3500	Micros; all types of application except games
BSD	7200	Micros and minis; business, education and home applications
MENU	28000	Micros, minis and mainframes; all type of applications

Table 1. Size and coverage

This source coverage implies, that MENU is the only candidate when shopping around for software to run on an IBM mainframe computer, or that BSD is ruled out when looking for information about scientific-engineering applications.

Currency may be judged by several aspects. The most straightforward is, the update frequency. BSD is updated quarterly, the other two databases are updated monthly, which is much more adequate considering the fact, that in a month several hundred programs are released, upgraded. Another way to get some impression about currency is to check how many records are available for relatively recent computers, like the ATARI ST family or the Amiga microcomputer. Yet another possibility which is also adequate to probe completeness is to check if description is available for some selected software. This is of course subjective, but it is understandable that for this paper seven gateway software were selected. The results are summarized in Table 2.

Criteria	Result		
	SOFT	BSD	MENU
Software for ATARI ST	3	2	117
Software for AMIGA	1	0	282
DIALOGLINK	yes	no	no
SEARCHMASTER	yes	yes	no
SCI-MATE	yes	no	yes
LAWSEARCH	no	yes	yes
LAWSEARCH PLUS	no	yes	yes
PRO-SEARCH	yes	no	yes

Table 2. Sample search results

3 SEARCHABILITY

Searching by many access points and by the combinations thereof is the greatest asset of online directories. All three databases offer an impressive amount of access points, still searchability falls short of expectations.

Essentially important access points are quite frequently entered in the records inaccurately, incompletely, inconsistently - or not at all, which make relevant records inaccessible. On the other hand, hardly needed data elements are used as access points.

It should be understood, that software is searched by professional intermediaries mostly by the following criteria:

- application type and functional capabilities
- operating system
- type of floppy disk on which the software is distributed
- memory requirements
- special hardware requirements
- price
- possibility to copy (for archival, back-up purposes)
- availability of demonstration version
- availability of toll-free telephone support

Fairly adequate searches can be made in all the three databases by application type and functional capabilities by using the descriptor, abstract and product description fields. MENU uses also a descriptor code by type of application to facilitate searching by this criteria.

Obviously, the abstracts and product descriptions are unable to provide all details, technical features, the potential buyer must consult product reviews and evaluations in trade journals for this. Therefore, it is very wellcome, that SOFT provides bibliographic details of software reviews in many records. (It is another question, that the structure of these bibliographies is very redundant and space wasting.) This feature was promised by BSD by the end of 1985, but it is still a promise.

Apart from functional capabilities, there are other, equally important technical characteristics of software, most prominently

- the operating system version under which the program can run
- the type of floppy disk on which the software is distributed
- the memory and special hardware required to run a program.

An application program may have several versions and editions requiring different operating systems, floppy disks, memory size and special hardware devices. This is well exemplified by the Wordperfect software family which has various editions among others for:

- a) MS-DOS 2.0 and PC-DOS 2.0 operating systems and their higher versions on 5.25 inch floppy disk with minimum 128Kbyte memory required
- b) Apple DOS 3.3 operating system on 5.25 inch floppy disk with minimum 64Kbyte memory required
- c) MS-DOS 3.2 operating system on 3.5 inch floppy disks with minimum 256Kbyte memory required (for the Data General One)

and in additions to the full-blown Version 4.1 it has a junior and a personal version for the PC-DOS 2.0 operating system. Furthermore, the PC-DOS 2.0 and the Apple DOS 3.3 editions differ in capabilities as much as Shakespeare's original Hamlet and its castrated version in the Simplified Reading Series.

All these would justify separate records for each versions, editions, but the database producers lump together all versions and editions in one record. In this particular case MENU comes nearest to the requirements by

distinguishing some of the releases and versions describing them in different records, but it also omits and improperly specifies the operating system required to run the program.

It is very frustrating that none of the directories specifies precisely, completely and consistently the operating system versions required to run a particular program. Overall MENU takes the greatest care in the precise specification of operating system and BSD is far the worst by not specifying at all the operating system in several hundred records! Specifying PC-DOS and MS-DOS is the same as specifying the title of a travel guide or a reference book without the year it relates to, or its language edition. Operating system names always require a version number

Specifying brand-names of hardware in all the three databases does not compensate for this deficiency. An IBM PC/XT may run PC-DOS 2.0, PC-DOS 2.2, PC-DOS 3.0, PC-DOS 3.1, a version of CP/M-86; an Apple IIc can run in addition to its native operating system PRODOS, also a version of CP/M-80, etc.

The typical microcomputer program is distributed on floppy disks; game programs are also distributed on cassettes and cartridges. It is therefore essential to know the distribution media and its size. The same program can be available on 8,5.25 and 3.5 inch disks for the same operating system version. MENU and SOFT provide this information reliably, BSD ignores it. This will be a serious problem with the proliferation of the IBM PC Convertible, which uses 3.5 inch disks and the PC-DOS 3.2 operating system, so most of the PC-programs - distributed at present on 5.25 inch disks - will be converted to 3.5 inch floppies. Information on the size of distribution disk will be essential in searching for software running under the PC-DOS 3.2 operating system, which is also available on the PC/XTs and PC/ATs with 5.25 inch drives.

The size of memory required to run a program is obviously critical. One who has the right version of the operating system, the adequate disk drive still may not be able to use a program requiring 512 Kbyte, if s/he has only 256Kbyte available. Both in SOFT and MENU this information is searchable and is supplied consistently, but in BSD its specification is rather the exception than the rule.

It goes without saying, that price is an important criterion in searching for software. This information is searchable in MENU and limitable in SOFT, but is available in only 60% of the records. In the DIALOG version of BSD price is not searchable, though appears in the majority of the records; in its BRS version limitation is possible by price.

There are more and more users who are willing to buy only non copy protected software, because of the inherent inconveniences of copy protection in using the software and making a back-up copy for safety purposes. It would be desired to be able to limit a search by this criterion, still it is hardly available in any of the databases, let alone limitable.

The same applies to demonstration version. More and more software producers market versions of their software with limited capabilities for testing purposes. These versions which cost between 10-50 dollars are excellent means to test a software before buying it, even if with limitations, e.g.: one cannot print, but only display results, or enter only a dozen of records, etc. Availability of demo disks should be included in the records and the database should be made searchable or limitable by this criterion. Practically none of the database include this information in the records, let alone index it from a dedicated field.

Users of software in the United States may be fortunate enough to have available toll-free number for asking help in case of a problem. This service is a very valuable asset, and availability of a no-charge hot line should be indicated in the records and the database made searchable by it. Only BSD cares to supply this information in the fields which describe the services available, though not consistently enough.

Consistently applying these information in the records would not add much overhead to the records and could replace other information which are always present, though their value - particularly as search criteria - is questionable.

Both BSD and SOFT offer the name of the contact person of the software producer as an access point, which is simply absurd. MENU also used this data as a index term, but dropped it a few months ago when reloading the database. Searching of software by state, city name, zip code, telephone area code of the manufacturer of the software is very unlikely, because most of the software is bought thru dealer networks rather than directly from the manufacturer, still one can search by all these data elements in both SOFT and BSD.

Searchability by processor type in SOFT may seem to be a sophisticated extra feature. In fact it is misleading rubbish. There is the same Motorola 68000 processor in the Macintosh, in the Amiga and in the Atari ST micro-computers but a program which runs on the Amiga will not run on the other two or vica versa just because they have the same processor. It is never ever a search criteria in shopping for application software. Let these examples suffice for irrelevant access points.

4 RECORD STRUCTURE

The record structure of SOFT comes nearest to the optimum. It is well laid out, easy to review, and contains few redundancies, like the field specifying where the software can be purchased, which contains no new useful information by stating that "directly or through dealers". As was said earlier, specification of the processor used by the software is irrelevant not only for searching but also for display. References to software reviews are annoyingly structured, contain many unnecessary fields and many of those fields contain only the explanatory label without any data.

BSD applies a fairly adequate record structure, but indicating the brand names of hardware in both the hardware and descriptor fields is just wasting space and makes more difficult to review the record at a glance.

The record structure of MENU is awfully redundant and is a disaster typographically. Enlisting the information related to memory requirements, disk size and price for all the hardware brand names and for all the operating systems under which the program can run is not only very superfluous, but makes the records very difficult to read. This is further aggravated by the slash and dash mania. Slashes are used not only to separate different data elements, but also in the descriptor field. The obsession with slashes and dashes is fairly well illustrated by the notation 5-1/4-inch and 3-1/2-inch instead of the much more common and simple 5.25 and 3.5 type notation.

The basic problem with record structure in all the three databases is, that usually all versions, all editions of a software are lumped together in one record, though these versions and editions differ in functional capabilities, in operating requirements, in price, in release date, etc. There should be an entry for each separate release, version and/or edition

of a software, just as there are different entries in a book catalog for the different editions of a book for various languages, media and publication years.

5 CONCLUSIONS

Directory type of software databases could well complement the abstracting/indexing databases with software coverage if they included all those data elements precisely and consistently which are typically used in searching for software, and if those data elements were adequately indexed. Unfortunately this is not the case in any of the software directories reviewed in this paper. Directory type of databases are reloaded rather than updated. In doing so, existing records are modified as opposed to abstracting/indexing databases where updating means simply adding new records. It is therefore hoped, that the content and structure of software directories can still be modified to reflect more appropriately the specifics of software searching.