

COMPUTER BIBLIOGRAPHIC SEARCHING: DOWNLOADING AND UPLOADING USING A MICROCOMPUTER

Joseph G. Wible, Librarian
Lane Medical Library, Stanford University

Downloading, the electronic capture and storage of data, can be used to provide computer bibliographic search results in electronic form instead of, or in addition to, the traditional paper copy. This method of data capture is becoming more desirable as library users begin organizing their personal files using microcomputer database managers. Downloading can also be used for eliminating duplicate citations in a multidatabase search. There is a general misconception that downloading is faster, and therefore cheaper, than directly producing a paper copy of the search results, but this is not the case. Printers are available which can print at speeds of 160 cps, while accurate transmission of data over phone lines is limited to 120 cps. Therefore, given the present technology, one can print as fast as one can download. On the other hand, search costs can be reduced by uploading, the sending of pre-stored data to the host computer.

Lane Medical Library is currently using a microcomputer for searches where uploading and/or downloading is desirable. The system configuration includes an IBM-XT, a Hayes Smartmodem 1200, and PC-Talk III communications software. Uploading is used routinely when the search requires extensive typing in an expensive database. Downloading is used less frequently because of its legal ramifications. Lane has collected "Downloading Policies" from the producers and vendors of medically related databases, and, at present, downloading is done only when the patron requests his search results on floppy disk, and when there are no restrictions placed on the database by the database producer or vendor.

Lane Medical Library is located in the Stanford University Medical Center and serves the hos-

pital, the medical school, and several University graduate programs. It is also used heavily by local pharmaceutical and biotechnology corporations in addition to being open to the general public. Online bibliographic searching comprises a major portion of the reference service at Lane with over 1000 searches being performed in 1983. This number is for full searches paid for by the user; it does not include the large number of ready reference searches done at no cost to the patron.

Over a year ago, the library was first approached by a physician interested in obtaining his search results on floppy disk. The library has had a few similar requests since then, but more frequently is asked for advice by patrons who wish to perform their own searches on their personal microcomputers. In response to these requests, the library purchased a microcomputer with a modem and communications software. It also began an aggressive end-user training program. The library has also taken on the role of consultant to the Medical Center community in regard to hardware and software necessary for accessing bibliographic databases.

Traditionally, online bibliographic searching has been done on a dumb printing terminal. Microcomputers can perform all the functions of a dumb printing terminal, and they have the advantage of being able to perform functions previously not available. Two such functions are uploading and downloading. Uploading can be defined as the transmission of previously created and stored data to the host computer. Uploading can be divided into two types. One is the transmission of an entire file which had been previously stored on a floppy or hard disk. This would typically be used for sending ILL requests or electronic mail.

The second type of uploading is the transmission of text, line by line, through the use of pre-programmed function keys. This type of uploading is more useful in online bibliographic searching where more flexibility is needed. Downloading can be defined as the capture and electronic storage of data received from the host computer. Both uploading and downloading are defined based on the perspective of the user. The same data which are uploaded by one computer can be downloaded by a second computer.

DOWNLOADING

Downloading offers a number of benefits. For example, it allows one to circumvent the problem resulting from a slow printer. A 1200 baud modem receives data at approximately 120 cps (characters per second). Letter quality printers or even some of the older dot matrix printers are not able to keep up with the incoming data and some of it would be lost. By downloading, the data can be received at high speeds and then printed later at a speed the printer can handle.

Downloading also allows the output to be edited before printing a final copy. The file can be loaded into a wordprocessing package where errors due to line noise can be corrected, irrelevant data and false drops can be erased; and citations can be merged and duplicates can be eliminated in the case of a multidatabase search. Notes and comments to the patron can also be added before printing.

Finally, downloading allows for the transfer of data directly into a microcomputer database manager without having to re-key the data. This aspect of downloading offers ex-

citing possibilities for the organization of personal reprint collections or, on a grander scale, the production of a personal database.

There are two common misconceptions concerning downloading. The first is that, by using a microcomputer to download, information can be received faster than by using a printing terminal. The rate limiting step is, in fact, not the printer but the phone lines. Normal phone lines can only handle 120 cps, and printing terminals are available which can keep up with this transmission rate just as well as a microcomputer.

The second misconception is one common among beginning end-users who think information has been downloaded just by creating the set of relevant citations. The user does not realize the citations must first scroll across the screen in order to be saved to disk.

UPLOADING

One advantage of uploading is it allows the typing of messages using a wordprocessor. This is advantageous for long blocks of text such as those used for ILL requests and other forms of electronic mail. However, for online bibliographic searching where the lines of text are short and flexibility is needed, this is probably of little value. The primary advantage of uploading for online searching is the ability to preprogram function keys with lines of text; each function key becomes a search statement. This allows the searcher to send individual lines of text by pressing one or two keys. By doing this one can cut down the connect time and thus reduce online costs significantly.

LEGAL CONSIDERATIONS

At Lane, uploading is done routinely while downloading is performed rarely. The primary reason downloading is not done is due to the legal problems surrounding downloading. A good discussion of the legalities of downloading can be found in the article by Thomas S. Warwick entitled "Large databases, small computers and fast modems...an attorney looks at the legal ramifications of downloading" in Online (1984) 8(4): 58-70. In December of 1983, Lane sent out requests for downloading policies of those database producers and vendors it most frequently uses. Lane asked the producer/vendor to interpret how their policy applied to the following four scenarios:

1) A librarian receives a search request from a patron. The librarian downloads a set of citations with abstracts from your database and from the Medline database. These two sets of citations are merged, the records sorted, duplicates are eliminated, and the format is standardized. A printout is generated for the patron, and the file is eliminated from the computer's memory within an hour after the downloading took place.

2) A librarian receives a search request from a patron. The patron requests his bibliography be delivered in machine readable form rather than on hard copy. The librarian performs the search and downloads the set of citations onto a floppy disk which is then given to the patron. The library does not keep either an electronic or a printed copy of the bibliography.

3) A physician (end-user) performs a search on your database and downloads a set of citations with abstracts. He then runs a program

against the downloaded file. This program generates reprint requests for each record and then adds the bibliographic records to his personal database manager. The database manager provides author and subject access to his extensive personal reprint collection.

4) A member of a research team (end-user) performs a search on your database and downloads a set of citations with abstracts. He then runs a program against the downloaded file. This program generates reprint requests for each record and then adds the bibliographic record to the laboratory's shared database. This in-house database is searched frequently by all of the members of the research team.

The letters received in response are kept on file at Lane. For the most part, database vendors deferred to the database producers. The database producers varied widely in their downloading policies (Fig. 1). For example, records from Chemical Abstracts can be downloaded for an annual fee ranging from \$50 to \$8,000 while there are no extra charges to download records from Psychological Abstracts.

FIGURE 1

DOWNLOADING POLICIES

BIOSIS	\$0.15 per record \$0.30 per record with abstract
CHEM ABS	\$50 - \$8,000 annual fee
EXCERPTA MEDICA	\$25 to allow temporary storage for printing
MEDLINE	permitted for personal use "multiple use subsets" available for a fee
PSYCH ABS	permitted for personal/organizational use

Database producers usually distinguish between records being downloaded for personal use versus those which will go into a database being used by a group of people.

Medline is produced by an agency of the Federal Government, and as such it can not be copyrighted. In order to avoid legal complications, Lane is at present only downloading records from the Medline database when searched through the National Library of Medicine.

PRACTICAL MATTERS

The configuration used at Lane for searching consists of an IBM-XT, a Hayes Smartmodem 1200, an Okidata 92 printer, with PC-Talk III communications software. The library owns a number of other communications software packages such as Crosstalk XVI, Smartcom II, and Kermit, but PC-Talk is used preferentially. It is very easy to use, does everything that is needed, and also happens to be the least expensive. It also does not cost anything to try since the author of the program encourages users to distribute copies to their friends. If the user likes the program, he/she is asked to send a \$35 donation to the author. In return for the money, the user receives a copy of the latest version of the program and is kept informed of future enhancements. Using the Alt-K command in PC-Talk, up to forty lines of text can be preprogrammed into the function keys for uploading.

Smartcom II is another communications package commonly used on microcomputers since it comes free with the internal Hayes Smartmodem 1200. Smartcom II does not come with the external version, however. Uploading using Smartcom is done using its Macro function. It also has a feature which allows the storage of logon sequences for each computer

accessed. It takes a lot of the hassles out of logon procedures. If the feature is used, one must consider security before including passwords in the sequence. Anyone with access to the system and a little knowledge about microcomputers could find the passwords.

Scimate is a software package marketed by Institute for Scientific Information. It comes in two parts, a communications package and a database manager. The communications package is designed to act as a user friendly interface for searching on Dialog, BRS, NLM, and ISI. Unfortunately the menu system is very cumbersome, and as a result, forces the user to waste large amounts of time going through menus while connected to expensive databases. The database manager does not currently provide indexes so when it searches for a term in the database, it searches linearly through the file. As a result response time is extremely long for locating records in databases of even moderate size. Scimate is not recommended. It does not provide for cost effective searching, and there are many database managers available which will give more value for their purchase price.

Most of the big name database managers on the market are not oriented toward bibliographic or textual data. Some examples of bibliographic database managers are Citation, Personal Bibliographic System, Notebook, and BioSuperfile/BITS. The latter is endorsed by BIOSIS and, based on demonstrations, seems a reasonable product for handling bibliographic records.

Finally, there is a new software package on the market called Insearch. This package is advertised as a user friendly system for searching on Dialog. In actuality, it is not very user friendly, but does provide a very cost effective way for searching Dialog data-

bases. It allows for preformulating a search strategy before going online. Once online, search statements can be revised while the computer is working. The search results are automatically placed into a memory buffer for offline editing and printing. Since it is expensive and only useful on the Dialog system, it can only be recommended for heavy Dialog users.

SUMMARY

Microcomputers are easily adapted for use in online bibliographic searching. As an alternative to the printing terminal, they provide the searcher with the additional capabilities of uploading and downloading. These two functions allow for more cost effective searching and for the development of new and better products.

Author's Note: Since the writing of this paper, new versions of **Scimate** and **Insearch** have been released. According to advertisements, **Scimate** now allows for preformulation of search strategies and **Insearch** can now be used on both **DIALOG** and **BRS**.

