

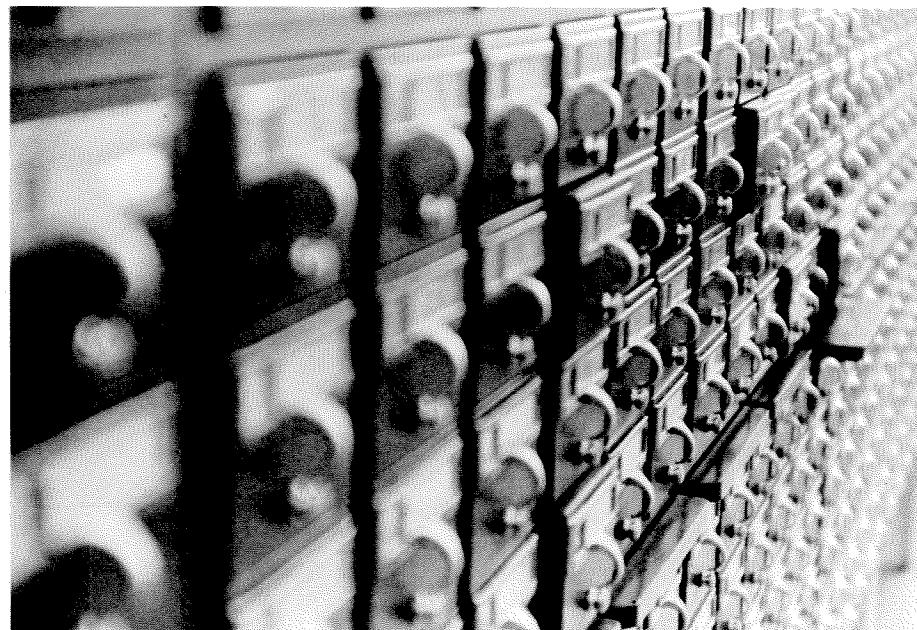
A Look at Technology and the Future of the Caltech Libraries

by Glenn L. Brudvig

INFORMATION TECHNOLOGY and libraries are both in a period of rapid change, with technology being a major driving force behind many of the changes now taking place in libraries. As the Caltech Libraries begin to move from a largely manual operation to a computer-based system, we need to have a clear idea of where technology is heading so that we can make today's decisions based on reasonable projections of future developments. The library needs to fit into an evolving information system which includes the use of personal computers on one end of the spectrum and large national systems on the other. It needs to maintain its role of being a link between a global range of information and a community of users, but increasingly that link will be an electronic one.

The use of personal or office computers is growing at rates that could not have been predicted just a few years ago. At present, they are being used in academic settings primarily for word processing (for composing and editing textual material), but also for statistical and computational computing, for data management, for production of charts and graphs, and for access to other computer systems. The personal computer is becoming the scholar's work station and a focal point from which many information and computer sources can be tapped, including local campus computers, national information systems, and library computer systems. It is also providing a new means for scholars to communicate with one another through electronic transmission of information over telephone lines.

The capabilities of personal computers continue to advance at an incredible rate. Within the not too distant future we can expect to see multifont page formatting, which integrates both text and graphics and which can generate publication-quality copy that can be transmitted electronically to a publisher. Joint authorship of publications by colleagues located in different parts of the country, or even of the world,



through the use of personal computers should become a common practice. Publications that are transmitted to publishers electronically can also be transmitted to colleagues or referees in the same manner. So, if writers write on line, editors edit on line, referees referee on line, and authors exchange preprints on line, can publishing on line be very far behind? Electronic publishing, whereby a paper copy would be produced only to meet a user's personal need, seems inevitable at some point in the future, possibly in 10 to 15 years, or maybe sooner for some types of publications. Although it may be very difficult for libraries to make plans for electronic publishing at present, or even to fathom how such systems may affect library services, we still need to be prepared to use them once they become available. The library systems that we develop today need to be planned with this eventuality in mind.

The personal computer will provide the potential for the scholar to gather information electronically from many sources, to build personal knowledge bases, and to construct a data base which will provide access to information now kept in file cabinets, note files, reference works, or reprint collections. Libraries should become an important source of electronic infor-

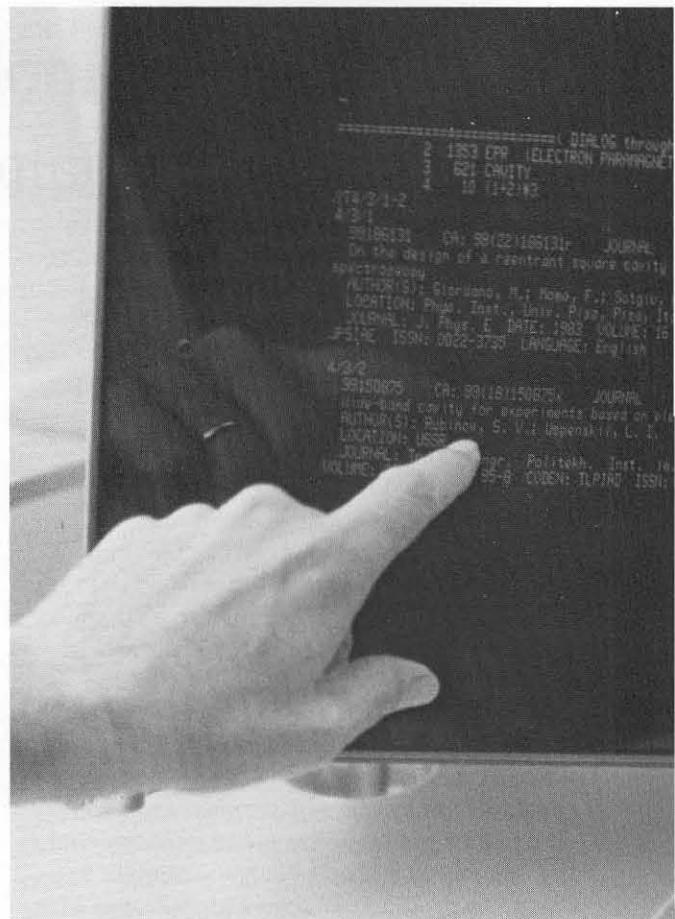


A year's worth of Chemical Abstracts, containing about 500,000 entries, fills three shelves. The information contained in this publication can also be retrieved on line at Millikan Library.

mation now and in the future. An important feature, therefore, of the initial plans for a computer-based library system for Caltech should be provisions for faculty members and students with personal computers to query the library's "card" catalog and indexes, to identify appropriate references, and to build bibliographies by transferring records electronically into their own files. A researcher's bibliographic files could also be automatically updated through a descriptive profile matched against the library's latest bibliographic records.

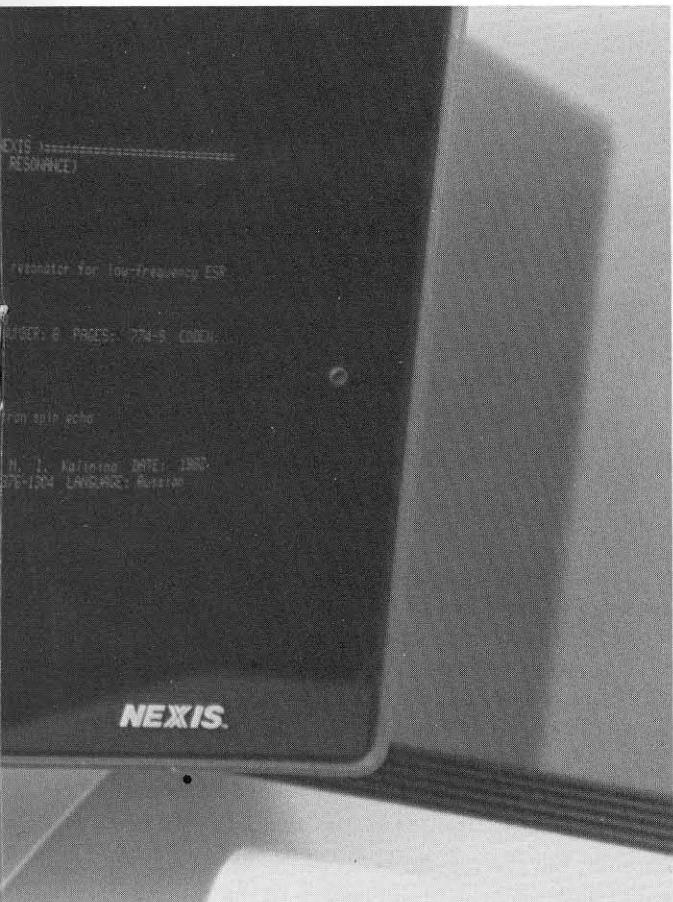
On the national scene, changes are taking place that will profoundly influence the kinds of services that libraries will provide in the future. What libraries will do in the future will depend to a very large extent on what the publishers do. When publishers shift from print-on-paper to electronic publishing, librarians will need to adjust their practices accordingly. Although this may be some years away, we are already beginning to see a shift from print to an electronic means of disseminating information in some areas. This shift has been taking place steadily in the publication of indexes and abstracts.

All of the indexing and abstracting services (for example, *Chemical Abstracts*, *Biological*



Abstracts, and *Reader's Guide to Periodical Literature*) are now produced by the computer. Most are accessible, or are becoming so, through on-line computer terminals. Although most indexing and abstracting services continue to produce printed versions of their products, some of the newer ones (those established after on-line searching became commonplace) are available only in an electronic format and have no print equivalents. Most libraries that use computer terminals for searching the literature also continue to purchase the printed indexes and use the computer only for the more complex searches that are impossible or difficult to do manually. There are a few libraries, however (new, specialized ones that have been established fairly recently), that have decided not to buy the printed indexes but to rely entirely on the computer for access to the literature. They have found it cheaper to do this than to purchase and store printed indexes. We are beginning to see a migration from print to electronic methods for information transfer in this specific area, but there are other examples as well.

A wide variety of data bases, in addition to the indexes and abstracts, are becoming available on line — for patents, statistics, biographical information, drug information, and so on.



Many of these are geared toward the commercial sector, but a large number are aimed at libraries as well. These data bases can be thought of as an on-line version of the reference book. Many of these newer, nonbibliographic data bases do not have a printed counterpart, and a few that once had them have dropped them.

Access to the full text of some publications is also becoming available through electronic means. The *Encyclopaedia Britannica*, the *Federal Register*, *Business Week*, *The New York Times*, and *The Washington Post*, to name a few, can all be retrieved in full text through a computer terminal. The 18 journals published by the American Chemical Society can also be retrieved on line in full text, as can the *Harvard Business Review*. The use of the computer to retrieve an article for one's general reading enjoyment is neither economical nor aesthetically acceptable, at least not at this time, but it is an excellent way to locate specific information within the text of an article that may be very difficult or time consuming to find in any other way. And for rapid access to material that is not immediately available, it can be a very valuable resource.

The rate of change from print to electronic

formats may be speeded up through new technologies, such as optical disks or more powerful software; but it could be slowed down by the costs of using such systems, by public acceptance, by inadequate copyright protection, or by a plethora of confusing and incompatible networks. Librarians need to stay abreast of these developments and evolve strategies for the future which can blend the new technologies together into a single system.

Because of the technological changes which have been taking place, the information resources of a library no longer need to be limited to the printed information in its collections. The library can move beyond its walls, encompassing an ever larger world of information which it can make available to its users, continuing its traditional role as the link between the information producer and the information consumer.

In planning an automation system for the Caltech Libraries, certain basic principles seem to be in order. First, as I have indicated, a library automation system needs to be a bridge between national information networks and personal computers, a bridge that can be used to transfer information to the individual user when it is needed. Second, it should be dedicated specifically to library operations, yet be able to interface with many different systems. Third, it should be an integrated system that will handle the full range of library and information retrieval functions. And fourth, it should be flexible so that it can continue to be expanded to take advantage of new technologies and to meet new demands that are placed upon it.

In planning for library automation, we need to begin with the ordinary "housekeeping" operations of the library — ordering books, handling the book fund accounting, checking in journals, and the activities related to getting a book onto the shelf. This is the process that will create an on-line catalog which is the index to the holdings of the library system. This process also involves a computer link to a national data base. For example, when a book has been recommended for purchase or is cataloged, the On-line Computer Library Center (OCLC) data base, which includes the cataloging records from the Library of Congress and 3,300 or so other libraries (including Caltech's), can be queried. If the record is found, it can be electronically transferred or copied into the library's computer file. This record can then be used to prepare a purchase order, an in-process record, and a catalog record. Very little typing

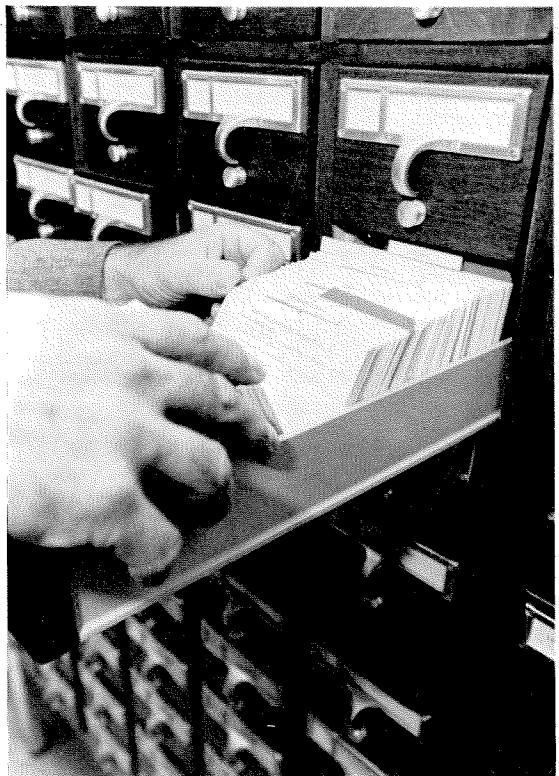
is done in the process, and most paper and manual files are eliminated.

The building of the library's on-line catalog also requires the conversion of existing cataloging records from 3-by-5 cards to machine-readable records. Once this was completed for the most used portion of the collection, we would add circulation control to the system. This would enable users of the system, through terminals located in the various units and libraries on campus, to determine not only whether an item was held by the library but also whether or not it was checked out.

By putting the card catalog on line, information on the resources held by all libraries on campus would be available at any point on campus. With collections divided between 10 floors in the Millikan Library and 10 departmental libraries, the problem of knowing where a particular item is located can be a time-consuming process. With an on-line catalog, users would also have more in-depth access to library holdings. They could search the catalog records in the usual way, by author, title, and subject, but also by the keyword in the title, by limiting a search to a publishing date, call number, or publisher, or by combining terms to either expand or limit a search.

The card catalog includes only the books held by the library. At Caltech, with its primary emphasis on science, it is the journal literature which is the most important source of information. Thus we need to improve our access to this resource as well. Before World War II, when labor was cheap and more plentiful and the number of journal publications was considerably less, some research libraries cataloged not only books but also the articles in the journals which it received. Their card catalogs thereby represented an index to the total resources held by the library. Now we have an opportunity to again do something very similar, but to do it electronically. Citations to the articles included in the journals currently received by the Caltech Libraries could be captured from national data bases, either by a monthly magnetic tape received on subscription or possibly by a computer-to-computer transfer of data.

With some 50 million journal articles currently accessible on line, this vast amount of information needs to be filtered so that we get only what is most relevant to the users on campus. If journals are selected carefully and quality is maintained, then the citations received by matching our current subscriptions against a large data base should reflect that quality. Those 50 million articles would still be



accessible on line for a charge for individual searches, but the regular users of the library would have access without charge to the smaller numbers of citations in current journals. Files of the most-cited articles could also be obtained in a similar manner.

The process of developing library automation moves step by step and takes time. With "housekeeping" as the first function to be automated, with an on-line catalog next, then circulation records, and finally a subset of the journal literature, a total automated library system can be built. At an appropriate point in the process, we would have to make the library computer system accessible to individuals with personal computers. A faculty member with a personal computer in his office, as was mentioned, should be able to search the library's data base and build bibliographies in his own files. He should also be able to request a book loan or a photocopy from the library, transmit a reference question or a literature search request, recommend that a new title be ordered, or request that he be notified when a book in circulation is returned. Materials would still need to be delivered or picked up, but at some future stage of development, a request for a journal article could result in the transmission of the full text on line. This is far beyond any current plans, but the transition from a request for a photocopy of an article to a request for the electronic transmission of its full text should



The familiar but cumbersome card catalog may eventually be on its way to extinction. At left, Charlotte Pavelko converts card catalog records into a machine-usable format.

be an easy step to make, at least in concept.

Even as it becomes possible to access an ever larger world of bibliographic information and more specific information hidden within that larger body of knowledge, something still is missing — methods for evaluating the published output. We can establish filters to screen the information from a global range of sources, such as matching our journal holdings against national data bases and obtaining files of most-cited articles, but that may not be enough. We may need additional methods for identifying the quality works in the literature. One method of doing this would be to have the computer keep track of use. This would identify the more heavily used, or popular, materials held by the library. Another approach would be to enable faculty members to indicate titles of high quality. It may even be practical to allow selected library users to comment on a book or article, giving the next user some indication of the value of the material being selected.

Once the library's bibliographic records were completely in the computer, it would then be possible for the librarian to enhance the record, to add information which would be useful to readers. Guides or pathways to the literature could be built into the data base. A library user could then determine what were the key texts in a particular field, the leading journals, the latest review articles, the special data bases, and so on.

Building the expertise of the reference librarian into the data base also becomes possible. The computer system could give instructions on how to use the system or the library, suggest search methods, provide directions to other sources of information outside of the library, or identify titles that need to be consulted to meet a particular information need. The possibilities of improving access to the literature increase tremendously with a computer-based library system. The development of the automated system for the library is therefore an open-ended process, as each new stage of development opens up new options for further development.

As we begin the development of an automation program for the Caltech Libraries, our immediate objectives are much more limited than the possibilities may allow. Our first steps will be to replace our manual operations with an automated system. To head up this effort we are now recruiting a systems development officer. We will be drawing up specific plans, establishing priorities, selecting software, purchasing equipment, and getting the project moving within the next several months. As we move ahead with this program, we intend to continue to take a long-range view of technology to help assure that our efforts will be compatible with the direction technology is headed and yet meet the immediate needs of the Caltech community. □