DOCUMENT RESUME

ED 284 594 IR 052 064

AUTHOR	Shaw, Ward
TITLE	Technology and Transformation in Academic Libraries.
PUB DATE	Jan 87
NOTE	17p.; In: Libraries and the Search for Academic Excellence. Proceedings of the Arden House Symposium (New York, New York, March 15-17, 1987); see IR 052 055.
PUB TYPE	Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	*Academic Libraries; Access to Information; Higher Education; *Information Dissemination; *Library Automation; Library Networks; Library Role; Library Technical Processes; *Online Catalogs; Online Searching; Technological Advancement; *Users (Information)
IDENTIFIERS	Colorado Alliance of Research Libraries

ABSTRACT

Academic library computing systems, which are among the most complex found in academic environments, now include external systems, such as online commercial search services and nationwide networks, and local systems that control and support internal operations. As librarians have realized the benefit of using computer systems to perform traditional tasks and the installation of systems has become relatively routine, library activities have shifted to encompass the operation of sophisticated computer systems, creating a conceptual gap between how libraries and librarians perceive themselves and how they are perceived by their campuses. With the introduction of the Online Public Access Catalog (OPAC), the gap has widened dramatically, as librarians have pioneered the creation of systems to deliver useful, interactive, computer-based services directly to end users. As the experience of the Colorado Alliance of Research Libraries illustrates, the ability of the OPAC to deliver information in real time to remote locations, access off-campus databases, provide multiple access points to information, and function as a general-purpose information utility has radically changed the model of libraries and their relationships to their clientele. Academic institutions must recognize these changes and ensure that they are implemented to provide maximum benefits to libraries, students, education, and research. (KM)

TECHNOLOGY AND TRANSFORMATION IN ACADEMIC LIBRARIES

by

Ward Shaw
Colorado Alliance of Research Libraries

Prepared for the Arden House Symposium
Libraries and the Search for Academic Excellence

January 1987

TECHNOLOGY AND TRANSFORMATION IN ACADEMIC LIBRARIES

by
Ward Shaw

This background paper is intended to provide a brief overview of the use of computer technology in academic libraries, a sample of some of the directions in which the use of technology is moving, and a description of a few of the emerging issues and opportunities these activities create. The subject is both large and complex, and this paper pretends neither depth nor comprehensiveness. Rather, it is meant to provide a point of reference and departure for the discussions of the Symposium.

One immediate problem with such a "broad brush" view is that it is necessarily an amalgamation. Some (not many) academic libraries make no direct use of computer technology at all. Others are involved in leading edge research and development far beyond what is discussed here. Either group may find the discussion irrelevent. However, the former should know what it has missed (or avoided), and the latter should know how far out they are. More importantly, to develop a common sense of direction and progression we need to know more about

the forest than any individual tree.

At the outset, we should observe that library computing is quite different from most other kinds of computing. First, it is characterized by literally enormous files, and the files are seldom used as a whole; rather an individual record within a several million record file is retrieved, a few characters modified, and the record replaced in a typical transaction. Second, there is much more reading of files than writing them, by perhaps 30 or 40 to 1. Third, there is very little numeric computation involved; in fact, we have been unable to identify any library "numbers" that are really numbers in a computation sense - instead, words and concepts are the entities of interest. Fourth, the use of library systems is almost always transaction oriented and in real time as compared to batch oriented with users prepared to wait. Fifth, library computing is almost always performed by people with little or no computer expertise, and often little interest in learning it. Sixth, library data structures are among the most complex ever encountered with typically dozens of fields and hundreds of subfields in a single record, all variable in length, format, and content. The document which describes the MARC standard for the construction of bibliographic records is more that 1000 pages long. Finally, users of library systems often approach the system with only a vague idea of what they want to do.

Successful library systems, therefore, are grossly optimized for moving information around, incorporate highly sophisticated file management techniques, are radically transaction oriented, require vast amounts of data storage, and very advanced user interface technology. They do not fit well in general purpose computers used for a wide range of applications, both because their consumption of specific kinds of resource is voracious, and because their design usually requires exploiting the characteristics of the hardware and operating systems at a fairly intimate level, which often has disastrous effects on other applications sharing the same computer. This explains why most library systems development has been on "stand alone" rather than shared machines. It also explains why library systems are among the most complex found in academic environments.

The use of computing technology in academic libraries divides into two basic areas - the use by libraries of external systems for access to external information and the operation by libraries of internal systems for control and support of internal operations. External systems are those owned and operated by organizations outside the control of the library from which the library purchases services; conversely, internal systems are those owned and operated by the library. There is a third category in which systems are operated by the university and used by the library, which creates a special set of

issues beyond the scope of this discussion. For convenience, these latter are considered internal systems here.

There are two main categories of external systems in widespread use. First, most academic libraries have accounts with and use one or more of the large online commercial search services such as Dialog, BRS, Mead Data, and the like. These systems, run by private companies for profit, acquire numbers of large files of mostly bibliographic and abstract information from various sources, and make them accessible to trained searchers through powerful retrieval software. Examples of such files are Chemical Abstracts, ERIC, Psychological Abstracts, and so on. Only occasionally do the end users of the information actually perform searches, as the software is powerful, complex, and expensive. Accordingly, libraries have trained specialists expert in search negotiation and execution who work with researchers, formulating and refining information requests, and operating the system to extract the appropriate data most efficiently. Costs are transaction sensitive, and usually include a basic subscription fee, hourly connect time charges, and a fee for each record retrieved. Typical costs for a given search can range from $10 or $15 up to several hundred dollars for the most complex and broad ranging requests. Few libraries can bear the costs of unlimited searching, and so usually ask the user to pay some or all of the expense. Although no research project is

complete without such a comprehensive review of the current literature as these systems provide, cost and complexity have tended to limit their broad use.

The other category of widespread use of external systems is for the support of the library's operations. Most academic libraries are members and users of one of the nationwide networks - OCLC, RLG, and WLN, for example - created by libraries to provide shared services. The basic idea behind these networks is resource sharing, and they have been most successful in two areas - first, as a source of catalog records, and second, as a location and messaging mechanism for interlibrary loan. The basic idea is that when any one library catalogs a book or other item, the record is stored in a large database accessible to all the members, who can use that record as the basis of their own cataloging. As the records are used, they are updated to indicate who owns the items described, and the files thus function as location tools. The scope of these nationwide networks is extraordinary. OCLC, for example, supports 7000 terminals in more than 5000 institutions. The data file contains 15,000,000 records. The system has been used to process more than 12,000,000 interlibrary loans. Individual institutions can, through access to such systems, enormously extend the resource they offer their clientele. And properly administered, they can save money over time by reducing the

rate of increase of processing costs. Like the commercial search systems, they are for the most part used by librarians either in their technical activities or on behalf of users, and only rarely by users themselves.

In the last decade, many academic libraries have begun to use local systems to support internal operations. Major advances in both hardware and software technology, and in the creation and promulgation of standards relating to hardware interoperability and data structures, have made it economically realistic and politically possible for individual institutions to engage in local systems development, and for an industry providing turnkey local systems to the library marketplace to emerge. Until the last three or four years, these systems were almost exclusively designed to support and manage internal housekeeping operations, such as circulation control, accounting, acquisitions, serials check in, and the like. These operations, when manual, consume surprising amounts of clerical effort, and even then don't work particularly well. First attempts at automating these functions did not work very well either, primarily because their complexity was underestimated. To the casual observer, circulation control looks very much like inventory control, and more systems than we care to admit failed because the designers did not understand the complex public services inherent in the special

functions libraries have created over time. Reserves, holds, recalls, fines, multiple locations, multiple media types with special loan policies, and so on, are all critical public services that circulation control systems must support, and may not be appreciated by system designers unacquainted with library practices. They are, however, appreciated and demanded by library users. But as systems and their design came more under control of the libraries themselves, the installation of systems to support these areas has become relatively routine and unarguably beneficial in all but the smallest situations. They do, however, represent major shifts in the kinds of activities engaged in by libraries, from the traditional "marking and parking" of books to the operation of sophisticated systems. This shift has and continues to create a conceptual gap between how libraries and librarians perceive themselves and how they are perceived by their campuses.

This gap has widened dramatically in the most recent three or four years with the introduction in libraries of Online Public Access Catalogs. All of the applications we have described to this point are systems used by librarians to do better what they have traditionally done. They have not changed the basic paradigm of librarianship, or the basic relationships of the library to its clientele and campus. Recently, however, that paradigm and those relationships have begun to

change rapidly and inevitably.

As the computer industry and its market has recognized that computers are fundamentally information machines and only incidentally computing machines, they have begun to design machines oriented towards real time transaction processing and huge file management and manipulations. At the same time, the idea of microcomputers has raised the possibility and to some extent the reality of extending to the general public access to information processing capacity not previously realistic. Librarians, who have always operated large transaction oriented dissemination and control systems, albeit manual, and who are professionally familiar with how people seek and use information, have not suprisingly been pioneers in creating systems to deliver useful interactive computer based services directly to end users.

Online Public Access Catalogs, replacing and extending manual card catalogs, have met with enormous user success, even in their infancy. The use of these systems is truly phenomenal. For example, 25,000 people each day use the CARL System, generating 1.6 million transactions as they interrogate and manipulate the various files we have available. We have nowhere near satisfied the demand. The only other kind of application with anything like that level of penetration

to the general public is automatic teller machines operated by banks, and those systems have little or no intellectual or pedagogical content. There are four characteristics of Online Public Access Catalogs (OPACs) which, taken together, radically change the paradigm or model of libraries and their relationships to their clientele.

First, one does not have to go to an online catalog to use it. In a manual environment, in order to use a library and its catalogs, the user has to go to the library. This is inconvenient, takes time and therefore must be planned in advance, and for almost all purposes restricts the library's utility to those users located in reasonable physical proximity to its facilities. In an automated environment, anyone anywhere with access to a telephone and a terminal or personal computer can use the system as effectively as someone in the library building. This single characteristic provides the institution with the potential capability of delivering one of its most important services anywhere, not just on campus. As telecommunications, telefacsimile, and storage capabilities advance, it will become possible to provide access to the source data and documents themselves as well as to index records describing those data. At CARL and in other places a number of experiments are already under way leading to this capability. This will create all kinds of opportunities, and problems, for academic libraries and their parent institutions that we are just beginning to

think about how to address. The most obvious opportunity is that just identified above - that is, the ability to deliver research information in real time to wherever it is needed or wanted, including, of course, off campus to business, industry, and homes. The most obvious problem is that of designing appropriate models to handle the proper compensations involved - users, universities, systems, libraries, publishers, authors, and so on - of which copyright is just the tip of the iceberg.

The second characteristic of OPACs important to note here is analagous to the first. Just as, at the user end, one does not have to go to the system to use it, likewise at the data end the systems can be used to access databases not necessarily resident on the campus, or on the local system. Several local systems, such as those at CARL, the University of California, and the University of Illinois, already contain data describing the resources of groups of institutions rather than only one, and make it very simple for users to examine the contents of multiple collections. And, considerable research and experimentation is occuring in the area of system interconnection. Again, the convergence of telecommunications and computer technology is the driving force here. A great deal of work is being done now in the development of standards and protocols for the interconnection and interoperability of computer systems, and much of the pioneering work

is going on in the library context. Research universities have for several years had access to Arpanet, connecting their academic computing centers, but most of the traffic on that network has been oriented toward access for advanced research to large computing capability. As the newer standards such as the Open Systems Interconnection Reference Model (OSI) are defined, promulgated, and implemented, it will become commonplace for all users of public access systems to routinely examine collections of institutions across the country for research information, and such interconnections will not be restricted to library databases, but will extend to all kinds of research information files. At CARL, for example, we support interconnections with several other library systems, and with several of the local campus computers so that anyone connected with those systems have access to us. We are also currently perfecting a link to the University of California Division of Library Automation's system, which will soon enable both groups of users to routinely access each others' files. Again, both the opportunities and problems are legion, and again, most of the opportunities relate to extended capabilities and most of the problems relate to economics.

The third characteristic is more of a "sleeper." Manual catalogs of any size are necessarily linear, organized usually by alphabet and according to headings or entries assigned by catalogers. Typically,

five or six access points are provided for any given bibliographic record. In computerized systems, however, the number of access points is limited only by the contents of the record itself, and those paths can be endlessly combined and related by users according to their individual requirements. Further, the system's messages, responses, and next choices can be based on the context of the given user interaction rather than predefined. The notion that computer systems are inherently de-humanizing is, therefore, exactly wrong. The best of them enable the transfer of control to the individual user, and encourage him to construct his own interaction with the research environment. The educational possibilities of this characteristic are just beginning to be explored, and we are just starting to take advantage of the feedback and heuristic capabilities inherent in the technology. Considerable research is in progress attempting to better understand the research process, and the mechanisms by which new information is created and enters the research infrastructure. When the interaction of personal computers and large search systems is considered, the design of enormously productive research work stations capable of supporting the collection and manipulation of large and diverse amounts of information, from many sources unconstrained by geography or time, toward the preparation of a variety of products ranging from publications to new learning, becomes a real and attainable goal. As this area is better understood, its extension to

all users is likely, and will clearly change the way almost all academic activities on campus (or off) proceed. What is critical to understand is that the process of informing becomes as much an element of design as is the database and its manipulation, and that the two are inextricably interrelated. The shift in the model here is dramatic - the library and its systems are suddenly active participants or even leaders in the educational process, rather than reactive service bureaus. Here also, there are opportunities and problems. The opportunity is nothing short of improving the quality of both learning and research both locally and broadly. One obvious problem is conceptual - that is, we do not yet understand either learning or research well enough to know much about how to approach the task. Another is ultimately political - as libraries become active rather than reactive, the transformation will be resisted and misunderstood.

The fourth characteristic is less esoteric, and is that we are discovering that other kinds of data files work well in the structures we have invented for our bibliographic files. Again using CARL as an example, we have several kinds of non-bibliographic information accessible to users of the system now, and are exploring a wide range of others. An almanac and a digest of current esoteric facts are currently available and generating considerable use. We are exploring files as diverse as a bible, complete with a concordance, a general

purpose encyclopaedia, the Official Airline Guide, a dynamic file of incoming airline passsengers and flights at Denver's airport, commercial and residential real estate advertising, regional statistical data, various indexes to newspapers and periodicals, financial data, classified advertising and more. This is not a wish list - in fact, we are not sure that we really want all of those - but each represents a real proposal before us right now. The point is that our user community perceives us as a general purpose information utility, and proposes to us that we do things that are considerably beyond our traditional conception of who we are and what we do. Once again, both opportunities and problems are present. One opportunity is that we might actually make money instead of consume it; and one problem is that we have no idea where we should draw a line to ensure that we are appropriately serving our primary purpose of bringing the best resource to the research and educational process.

It is clear that the application of technology in libraries has been and continues to be extraordinarily robust. Contrary to popular myth, libraries have collectively and enthusiastically embraced new technology when its use is appropriate, and created the research projects and experiments necessary to lead the transitions. It is also clear that the same process is causing a radical redefinition of the academic library, its role in the educational process, and its

relationship to the rest of the institution. These changes are inexorable - the benefits are too great and the economics too compelling to resist. But they create confusion in various institutional perceptions that must soon begin to be addressed. Academic libraries have traditionally been the arena in which the leaders of our profession created techniques and mechanisms to develop, teach, and support the relationship of information and ideas. It is my observation that, for whatever reason, that leadership has shifted to public libraries, which seem to have much more flexibly redefined their roles and activities, and are more and more frequently conducting the forward looking experiments. It is important that academic institutions begin to recognize that there is a profound set of changes under way, and quickly direct attention and resource to ensure that the changes are designed and implemented for maximum benefit to libraries, students, and to education and research.