DOCUMENT RESUME

ED 060 913 LI 003 583

AUTHOR	Ferguson, Douglas
TITLE	The Library, the Researcher, and Computerized Information at Stanford University; A Report to the Director of Libraries.
INSTITUTION	Stanford Univ., Calif.
PUB DATE	Oct 71
NOTE	23p.; (7 References)
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	*Computers; Information Needs; *Information Systems; *Library Services; *University Libraries
IDENTIFIERS	*Library Role; Stanford University

ABSTRACT

This report is intended to aid the library at Stanford University in determining: (1) How it should respond to faculty and student need to know about computer-based information systems; (2) If the library has a role in supporting awareness of, and access to information about computerized files on campus; and (3) If the library has a role in planning services centering around large data bases that will likely soon be available on campus. The author concludes that: (1) There is evidence of a need for information about computer-based information systems among faculty and students; (2) There is a sizable number of computerized files or data collections on the campus; (3) The library is the appropriate place to respond to the need for information about computer-based information systems; (4) The library has unique experience and skills that would be valuable in promoting awareness of and access to computerized files; and (5) It is to be university's advantage for the library to maintain active involvement with the groups that are developing a variety of computer-based information services on campus. It is recommended that the library augment, and actively support, services for answering requests about computer-based information systems. (Author/SJ)

THE LIBRARY, THE RESEARCHER, AND COMPUTERIZED INFORMATION AT STANFORD UNIVERSITY

A Report to the Director of Libraries

Douglas Ferguson
SPIRES/BALLOTS Project

October 1971

Table of Contents

THE AIM OF THIS REPORT

This report is intended to aid the library in determing: (1) How it should respond to faculty and student need to know about computer-based information systems, (2) If the library has a role in supporting awareness of, and access to information about computerized files on campus, and (3) If the library has a role in planning services centering around large, generally useful, data bases that will likely be available on campus in the next few years. A word about terminology. For purposes of this report, computer-based information systems will include both literature services such as Engineering Index and non-literature services of the descriptive data type as supplied, for example, by the Roper Center. Collectively, these will be referred to as computer-based information systems. Information stored on punched cards, tapes or disks is usually called machine-readable data but in this report the phrase "computerized information" is used on the analogy of printed information or recorded information. Files and data bases are collections of computerized information. In this report a file refers to a relatively small collection usually created by and for a few specialists, such as a file of questionnaire data. A data base refers to a somewhat larger collection often produced and maintained for many sizeable groups of users by a central organization, such as the MARC data base.

CONCLUSIONS AND RECOMMENDATIONS

Conclusions

(1) There is evidence of a need for information about computer-based information systems among faculty and students at Stanford. (2) There is a sizable number of computerized files or data collections in various places on the campus. (3) The library is the appropriate campus organization to respond to the need for information about computer-based information systems. (4) The library has unique experience and skills that would be valuable in promoting awareness of and access to computerized files on campus. (5) It is to the university's advantage for the library to maintain active involvement with the groups that are developing a variety of computer-based information services on campus.

RECOMMENDATIONS

Policy Recommendations

(1) Augmented library servicing of requests about computer-based information systems should be implemented as soon as possible. (2) Providing awareness of and access to computerized information at Stanford should be actively supported by the library. (3) The library should maintain liason with groups planning to acquire and maintain large data bases of generally useful information.

Implementation Recommendations

(1) Senior Librarians, especially branch librarians, should be consulted and encouraged to participate in shaping the mode of servicing requests for information on computer-based information services. (2) Any service that is established should be studied after six months of full operation to evaluate its usefulness, possible discontinuance, or ways to improve it. (3) A library administrator of at least department head rank should be designated by the Director of Libraries as the official liason with groups studying existing computerized data on campus and planning for computer-based information systems on campus. (4) A central file and services should be established for material on all major computer-based information systems.

BACKGROUND CONSIDERATIONS

In 1968 a group of faculty proposed the establishment of a data archive at Stanford. Because external and internal funding was not available, no concrete action was taken on this proposal. As a result of continued faculty interest, in 1970 the library recruited a person to work part-time on developing what was known as a data tape information service. A person was hired in January 1971 and because of that person's untimely death a few months later, the planning was assigned to me in April 1971. In

parallel with these developments the library has actively been engaged in project BALLOTS which is planning to implement a sizable data base (i.e. MARC) in the second quarter of 1971 if adequate funding is available. In addition, a research group in the Institute for Communications Research (Jesse Caton and Colin Mick) are developing an interactive statistical package called INTERSTAT which could be used with large and small computerized files. The SPIRES group is working closely with the BALLOTS group in developing software for the search, retrieval, and update of large literature and non-literature data bases.

Meanwhile, other universities have actively worked to meet their faculty and student needs for computer-based information services in a number of different ways. UCLA, under sizable grants from the National Science Foundation, is providing a data base maintenance and search service within the library using the library's dedicated computer. The University of Florida at Gainsville is providing a similar service on the campus computer but the library is the service point for users on a state university network. Reference librarians code all searches of the data bases (e.g. MARC or ERIC) and funnel the search results back to users whether they are on campus or at one of the other network schools. Cornell has established an experimental "Information Systems Clearinghouse" in its library staffed by one librarian. This office maintains up-to-date material on computer-based information systems and actively seeks out

researchers who might benefit from using them. These are three different examples and others could be added.

During August and September I conducted a census data users survey to determine the need for computerized census data on campus. A report on this survey and related matters was submitted to the University Libraries and Computation Center on October 6th. As a result of strong impetus from the University Libraries it appears likely that census data will be acquired and maintained at Stanford as the first large, widely used, research data base. The survey clearly established the existence of many small files in various locations on campus including the Campus Facility, ACME, department computer rooms, faculty offices, research laboratories, and possibly even homes. The Campus Facility provides tape storage services for its users and has an inventory type control of the tapes. It is not now possible for the staff to provide access (such as an index) to the contents of the tapes. Not only is this a large task but it is by no means certain that tape "owners" want the contents of their tapes disclosed publically.

THE LIBRARY AND THE COMPUTER CENTER: A PERSPECTIVE

The library is knowledge recorded on various media and skilled people who help make this knowledge accessible. A computer center is a powerful machine, programs that control its

operation, computerized information on various media, and a skilled staff of computer specialists. The computer and its related equipment is the major physical investment of the center just as the book collection is the major physical investment of the library. Each organization has a physical store, equipment, service points and a staff. Looked at in this way libraries and computer centers possess similarities as well as differences.

Both libraries and computer centers store large quantities of information and depend on trained people for their effective use. A computer center staff is more technology oriented than the staff of a library and a library staff has more public service positions than a computer center staff. A computer center usually has only one service point and the need for accessibility is met by placing computer terminals at various sites away from the computer center, or occassionally by setting up another computer center. Libraries invariably have many service points through a branch library system. Retrieval services performed by a computer center are mainly dependent on the computer while library retrieval services are mainly dependent on reference librarians.

The situation is changing somewhat so that the differences between libraries and computer centers are more of degree than of kind. In general libraries are using computer services and computer centers are becoming more oriented toward larger user

groups most of whom have little technical background. Printed indexes and bibliographies in libraries are often produced by computer. Computer centers often have sizeable collections of documents that explain the use of equipment, programs, programming languages, etc. It is increasingly difficult to distinguish between the public service role of a computer user consultant and the public service role of a reference librarian, though neither can perform the others work without appropriate training. Libraries use computer services in automation programs and computer center personnel depend more and more on a strong library of printed material, especially in a university setting.

Both the computer professional and the library professional will find this discussion ignores details and differences which each considers important. It is essential, though, to see the library and the computer center as two service organizations that are becoming more not less inter-related. They are moving closer toward a partnership characterized by cooperation and a mutual complementarity of services. Indeed the time is not far off when the university will consider them as an indistinguishable common resource of information and service supporting teaching and research.

ISSUES AND ALTERNATIVES FOR THE UNIVERSITY LIBRARIES

The library must take more than judicial notice of the

growing role of computerized information and computer-based information systems in the work of scholar and student. The reasoning behind this statement is as follows. There is a growing number of computer-based information systems with an increasing body of users. Research is a major activity at Stanford and these services support research with retrospective literature searches, on-going current awareness products, and individual SDI services. The high quality of computer services at Stanford has increased the acceptance and use of computer services among both students and faculty. Bibliographical control of information about computer-based information systems is not well developed either nationally or at Stanford. Material about these services is scattered throughout Stanford and much of it is probably not up-to-date. Once inforation has been computerized it can be re-used and re-analyzed at a low incremental cost IF IT'S EXISTENCE IS KNOWN TO POTENTIAL USERS. This lack of bibliographic control and the hidden quality of computerized files at Stanford (and also at nearby schools and organizations), wastes the time and money of Stanford faculty and students. The library has the knowledge, the skills, the experience and the dissemination system to improve this situation. Furthermore, an active role with respect to computer-based information systems and the scholar would put the library in closer touch with research activity and increase its ability to contribute to planning for campus computer-based information systems.

If this reasoning is accepted the following are the issues to be faced by the library in determining its service role with respect to computer-based information systems: what services should be provided, how should the services be provided, and at what cost? Another issue is immediate services and future services. The library's role vis-a-vis computerized files at Stanford will be discussed following the discussion of computer-based information systems.

LIBRARY SERVICE AND COMPUTER-BASED INFORMATION SYSTEMS

Library service directed toward improving access to and use of computer-based information systems, has several components. Each successive component increases the level of service, and also the cost. Each successive component combined with previous components is an alternative service configuration that the library can select. Here are the components of library service in this area.

(1) Reference service from generally available printed sources now in the collection. This represents no change in the present level of library service.

(2) Reference service from specially created and maintained manual files of material on each major system. This is an

additional acquisition effort (i.e. letter writing) and file maintainance task.

(3) Announcement or dissemination services. This may involve a list of profiles of each system and/or announcements in campus publications. This adds an analysis and summarization task.

(4) Direct assistance in preparing a service request to a system. This adds a specialized reference task involving the use of special request forms and thesauri, (e.g. lists of descriptors or index terms).

(5) Sharing the cost of using a system with the professor or student. This adds funding, accounting and fund allocation tasks.

(6) Assistance in formulating and coding search queries to locally maintained data bases. This adds a computer training task.

A service configuration consisting of components two and three requires 1/2 FTE librarian plus cost of mailing (both letters to the systems and on-campus dissemination) and supplies. A service configuration consisting of components two through four requires 1 FTE librarian, assuming a central service point. Purchase of thesauri might be necessary. The cost of a service

configuration involving components five and six cannot be estimated at this time. I recommend that a service configuration consisting of reference and announcement service be started within the next three months, A direct assistance component should be added after evaluating The first six months of service. A decision on expanding or remaining at the first implemented level should be made within one year.

Suggested tasks involved in implementing the first service configuration are the following.

(1) Select the major computer-based information systems about which material needs to be gathered.

(2) Determine what information exists on them in the several libraries at Stanford.

(3) Make a copy of up-to-date material on campus for the central file and send for current material on the remaining systems.

(4) Set up a central file of material arranged by the name of the organization supplying the service.

(5) Prepare a profile of each system describing the subject areas covered, services provided, how you send in a service request and the costs. This should be brief because details are supplied by

the material in the folder.

(6) Prepare a list of profiles by subject area listing interdisciplinary services under more than one heading as required.

(7) Send this list to all academic departments for posting and to all campus libraries. Issue a revised version no more than twice a year. For services that specialize in one subject area for which there is a library on campus, copies of material in the main file should be sent to the library specializing in that subject.

Here are some suggestions for detailed implementation. As a selection tool to accomplish task one above, begin with the DIRECTORY OF COMPUTERIZED INFORMATION IN SCIENCE AND TECHNOLOGY. Since this has a comprehensive subject index, it can also serve as a subject index to the central file if appropriately marked. The limitations of this work are that it does not contain information on current costs and it has no sample request forms or samples of the printed products or outputs. The library now subscribes to two copies of this work, one in the Engineering Library and one in the Computer Science Library. Since the annual subscription rate is $175.00 the library may wish to evaluate the need for the copy in the Computer Science Library. My personal observation is that it is not necessary and I think the Computer Science Librarian would concur.

To reach most researchers on campus consider contributing a regular section not more than two paragraphs long "Information Services for Research" to the RESEARCH INFORMATION CENTER BULLETIN. Highlight one or two services each issue, describing how it can be useful in state-of the-art surveys, locating parallel research projects, and keeping up with current research in related fields. This newsletter is issued every five weeks by the Research Information Center in the Provost's office and goes to all faculty and all staff researchers. Make sure that all librarians (including those in the Coordinate Libraries) know of the central file so they can use it and refer patrons to it.

Whoever maintains the central file should also be familiar with systems and organizations specializing in non- literature or data services, e.g. National Opinion Research Center in Chicago, the InterUniversity Consortium for Political Research in Ann Arbor, the Roper Public Opinion Research Center in Williamstown Mass. and of course the Survey Research Center in Berkeley. The important thing initially is to focus on the major services and get out up-to-date information on them. Comprehensiveness will come as the person(s) assigned gain experience. Contrary to what one would expect, researchers often do NOT know of the major computer-based information services in their fields. Less than two months ago the coordinator of a multi-million dollar project in the bio-medical sciences contacted me and he was not at all

familiar with MEDLARS. Several researchers who contacted me in my work with the Research Information Center appreciated knowing about computer-based information systems that would give them literature searching assistance as part of their research proposal preparation work.

PROMOTING AWARENESS OF COMPUTERIZED FILES AT STANFORD

The multitude of relatively small computerized files on campus have initial interest only to their creators and possibly to a few related researchers. Some files may have been produced at considerable expense and effort and might be considered proprietary until reports or articles have been issued. These are significant limiting conditions on the creation of a campus union (i.e. location) list of computerized files. Against these considerations must be balanced the fact that knowing about a data collection on campus might save one researcher hundreds (even thousands) of dollars in time and out-of-pocket expense that he might otherwise spend in replicating the data collection. Awareness of an existing file might give a graduate student or another researcher ideas for re-analysis of the data to extend or refine findings of previous studies based on the data. In promoting awareness of computerized files on campus the library would be fulfilling its person-to-person switching function by putting researchers in contact with one another to the mutual advantage of all. The research community at Stanford is large and

geographically dispersed. Anyone who has worked with interdisciplinary research groups at Stanford can attest to the frequency with which someone elses ongoing research dovetails with ones own after each has apparently been going on its own for some time.

How can a start be made in providing access to files on campus? You will recall that the INTERSTAT development group in the Communications Department is working intensively on an interactive statistical system that can be used to analyze many different types of files. As part of this work a preliminary survey of files on campus may be undertaken in the next few months. This could be the start of a campus listing of computerized files. Included in this survey should be questions that elicit the size, contents, ownership, location, and characteristics of the file and especially the conditions under which the owner is willing to share it. From the survey a list "Computerized Research Files at Stanford" could be prepared and distributed to all academic departments and libraries. Through notices in campus publications such as the STANFORD COMPUTATION CENTER BULLETIN and the RESEARCH INFORMATION CENTER BULLETIN, researchers could be encouraged to voluntarily list their files. This list could be exchanged with similar lists that may be available at Berkeley. It could also be distributed to libraries and computer centers at nearby colleges and universities, particularly those that are part of the various computer networks

related to Stanford. Other schools might then be encouraged to conduct similar surveys an in time a regional union list of computerized files might develop. Side effects of this work could include cooperative work between researchers at Stanford and at other schools and additional use of Stanford computer services.

THE LIBRARY AND UNIVERSITY SUPPORTED DATA BASES

The university library has a continuing responsibility to maintain and make available the printed resources of the university. As more information becomes available in computerized form faculty and students will use this along with the published literature. There are similarities between collecting printed and collecting computerized information. Both require money, acquistion guidelines, and updating procedures. They require some means of access or awareness to bring them to the attention of potential users as well as trained specialists to help people use them. There are also differences between a computerized information collection and a printed information collection. Individual computerized files are not normally made available widely and cheaply as are printed works. Computerized information is either created by a specialized research project in which case it is in private hands or at best in one of the major data archives or it is created by a commercial, governmental, or professional organization and access services are sold to users. In both cases there is no wide dissemination to free public service organizations such as libraries.

Data bases are, however, growing in number, organizations are increasingly selling the data bases as well as services, and data bases are being created that are useful to many disciplines. Wholesaler-retailer relationships are developing between central services and dispersed organizations. For example, ERIC or Chemical Abstracts produce tapes and companies or educational institutions buy them and provide tailored services to their local clientele. Such arrangements are not unknown to the library world, extending back to centralized cataloging and up to the Library of Congress MARC service and the National Library of Medicine's regional MEDLARS centers.

What is the role of the University Libraries as Stanford looks toward augmenting its computerized information resources? Purchase of a data base is a major university expenditure of the same order and importance as the purchase of a complete journal set, a subscription to an indexing or abstracting service, or the purchase of a private library, and in addition there are ongoing expenses and servicing requirements. The University Libraries has a role in promoting user involvement, in promoting user communication, and in providing supporting referral service. I suggest that the library and the computer center actively work to create a faculty, student, and staff user group on data base services. This group should develop a data base acquistion policy that sets up criteria for major purchases and makes

recommendations for the adequate servicing of data bases that are aquired. This group could also be a source of feedback on the usefulness of the services provided. The library should use its staff and branch library system as a means of making the universtiy community aware of the data bases and services. A formal library staff liason with the computer center and groups developing data base services (such as INTERSTAT and SPIRES) should keep the library staff aware of developments both informally and through periodic reports in the STANFORD LIBRARY BULLETIN. In this way the library staff can provide comprehensive reference services to printed and computerized information. With the purchase of the census data base in the near future, the library has the opportunity to pilot test this suggested role by continuing to be involved with work on this significant data base.

CONCLUDING COMMENTS

The recommendations of this report assume no major reorientation of library services, no major additional staff, and no major training effort. The recommendations suggest a response to a demonstrated need for which the library is well prepared. There is now an opportunity for librarians in various departments and in the Coordinate Libraries to work together with the Computation Center staff, as has been done in the past, to improve information services to the Stanford community.

NOTES

Current material on several of the major computer-based information systems has been collected and is in the SPIRES/BALLOTS documentation files.

THE DIRECTORY OF COMPUTERIZED INFORMATION IN SCIENCE AND TECHNOLOGY referred to in the text is in the Engineering and in the Computer Science Library (Z699.5 T4D5f). A copy that has been manually updated from the Computer Science copy is in the SPIRES/BALLOTS Documentation office.

The arguments pro and con for direct servicing of data bases by the university library are extensively discussed in Hayes, R.M MECHANIZED INFORMATION SERVICES IN THE UNIVERSITY RESEARCH LIBRARY. Phase I, Final Report (parts 1-13) UCLA, December 1967. Microfiche copies are in the Documentation file "UCLA--Center for Information Services".

A somewhat out-of-date but still useful work is the ALA GUIDE TO THE SELECTION OF COMPUTER-BASED SCIENCE AND TECHNOLOGY REFERENCE SERVICES. 1969. The Main Library Reference Service has it on order. A work of about the same date is the STUDY OF SCIENTIFIC AND TECHNICAL DATA ACTIVITIES IN THE US. Science Communications Inc. December 1968. AD 670 606 AD 670 607 available in the UC

Berkeley Government Documents library. A recent and useful work is Carroll, Kenneth D. SURVEY OF SCIENTIFIC-TECHNICAL TAPE SERVICES. American Institute of Physics. September 1970. PB 196 154. The microfiche of this is in my file of notes which is with Mr. Epstein. This file also contains miscellaneous correspondence and file memos including those on the Florida and Cornell approaches referred to in the report.

Computerized information in the Humanities is reviewed in Bowles, Edmund A. "Computerized Research in the Humanities: A Survey", American Council of Learned Societies Newletter Special Supplement. June 1968. pp1-49. An earlier work is Carlson, Gary LITERARY WORKS IN MACHINE-READABLE FORM. Computer Research Center Brigham Young University. July 1965 and updated in the January 1967 issue of COMPUTERS IN THE HUMANITIES. Both works are available at Stanford.

A copy of the inventory printout prepared by the now defunct Council of Social Science Data Archives dated April 1970 is in the above mentioned Documentation files. Prof. Carl Beck of the University Center for International Studies at the University of Pittsburg is the heir to this former NSF project. The inventory has not been maintained.

The Resources and Technical Services Division Committee on Descriptive Cataloging has a subcommittee on Rules for Cataloging Machine-Readable Data Files. John D. Byrum Jr. at Princeton's Library is its chairman. I wrote him in June but never received a reply. The subcommittee may be inactive.

DATE: October 25, 1971

TO : David C. Weber

FROM : Douglas Ferguson

SUBJECT: Report on the Library and Computerized Information

Attached is a report on what has heretofore been called the "Data Tape Information Service." I consider it to be a discussion paper since it is not as comprehensive or as detailed as I would have preferred. I tried to make it as straightforward as possible keeping in mind a possible audience in the University Libraries and in the Computation Center. The contents are my own responsibility and I have not distributed it beyond Mr. Epstein my immediate supervisor, Mr. Borgeson and Professor Parker in his role as your Special Assistant for Computer Services. I would think an expanded distribution list might include Allen Veaner, Susan Kolasa and Peter Nycum. Others such as Jack Pooler and the heads of the Branch libraries and the Coordinate Libraries would, I feel sure, be able to comment knowledgeably. I hope the report is useful to you and your staff in developing services in relation to computerized information. It was a valuable experience preparing it.

cc: Earl Borgeson
Hank Epstein
Edwin Parker