DOCUMENT RESUME

ED 253 231 IR 050 994

AUTHOR Aveney, Brian, Ed.
TITLE Online Catalog Design Issues A Series of Discussions. Report of a conference sponsored by the Council on Library Resources (Baltimore, Maryland, September 21-23, 1983).
INSTITUTION Council on Library Resources, Inc., Washington, D.C.
PUB DATE Jul 84
NOTE 250p.; Part of the Bibliographic Service Development Program. For related documents, see IR 050 993-996.
PUB TYPE Collected Works - Conference Proceedings (021) -- Reports - Descriptive (141)

EDRS PRICE MF01/PC10 Plus Postage.
DESCRIPTORS Academic Libraries; Cataloging; Computer Science; Guidelines; *Information Retrieval; *Information Storage; *Information Systems; Library Automation; *Library Catalogs; Library Networks; *Online Systems; Programing; Research Libraries; Search Strategies; *Systems Development
IDENTIFIERS Council on Library Resources; *Online Catalogs

ABSTRACT

Developed from presentations given at a 3-day invitational meeting of 31 leading online catalog designers from both public and private sectors in North America, these papers provide information about online catalogs--their present form, their use, and the questions that need to be considered in their future refinement. All of the papers in this volume have been edited from transcripts or orignal texts. In a few cases, papers have been extensively reworked. The document comprises: (1) "Data Structures and Resource Consumption" (John R. Schroeder and Jessie J. Herr); (2) "Search Retrieval Options" (James F. Corey); (3) "User Feedback in the Design Process" (Charles Hildreth); (4) "Screen Layouts and Displays" (Joseph R. Matthews); (5) "Command Languages and Codes" (Michael Monahan); (6) "Online User Prompts and Aids" (Clay Burrows); (7) "Linking of Systems" (Ray DeBuse); (8) "Telecommunications Considerations for Online Catalogs" (Edwin Brownrigg); and (9) "Summary Questions and Discussion" (Brian Aveney). An agenda of the meeting and a list of participants are appended. (THC)

ONLINE CATALOG DESIGN ISSUES

A Series of Discussions

Report of a conference sponsored by
The Council on Library Resources

at the

Holiday Inn - Inner Harbor
Baltimore, Maryland

September 21-23, 1983

Edited by

Brian Aveney

Bibliographic Service Development Program

Council on Library Resources, Inc.
1785 Massachusetts Avenue, N.W.
Washington, D.C. 20036

July 1984

The conference described in this report was funded as part of the Council on Library Resources' Bibliographic Service Development Program, a program administered by CLR and funded by several parties: the Carnegie Corporation, the Commonwealth Fund, the Ford Foundation, the William and Flora Hewlett Foundation, the Lilly Endowment, the Andrew W. Mellon Foundation, the Alfred P. Sloan Foundation, and the National Endowment for the Humanities.

Library of Congress Cataloging in Publication Data
Main entry under title:

Online catalog design issues.

"July 1983."
1. Catalogs, On-line--Congresses. 2. Machine-readable bibliographic data--Congresses. 3. Cataloging--Data processing--Congresses. 4. On-line bibliographic searching--Congresses. 5. Library information networks--Congresses. 6. Information storage and retrieval systems--Congresses. 7. Libraries--Automation--Congresses. I. Aveney, Brian, 1940-
Z699.A10542 1984 025.3'028'54 84-15530

ONLINE CATALOG DESIGN ISSUES

Contents

Preface

The content of this publication provides much information about online catalogs--their present form, their use, and the questions that need to be considered in their future refinement. The meeting is reported in some detail because interest in the topic is high among many more individuals than the thirty-one who took part.

By bringing together systems designers from libraries and commercial organizations serving libraries, CLR sought to provide a setting that would encourage coordination of effort and development of reasonable guidelines for the improvement of catalogs that might influence design work and, in the end, benefit catalog users.

Once again, CLR can express the thanks of the library community to the authors and participants. They worked together to help shape the future of online catalogs, which are one of the most visible signs that the library of the future is really here.

Warren J. Haas
July 1984

I. INTRODUCTION

The papers in this volume developed from presentations given at a three-day invitational meeting of 31 leading online catalog designers from both public and private sectors in North America. American designers A. Stratton and Caryl McAllister, architects of IBM's DOBIS, located in West Germany this last decade, were also able to join the group.

The sessions consisted of a series of challenge papers followed by extensive discussions. All of the papers in this volume have been edited from transcripts or original texts. In a few cases, papers have been extensively reworked.

In a paper updated from the session presentation by John Schroeder, Director of Research for the Research Libraries Group, he and Jessie Herr, former Manager of Database Systems at RLG, discuss the system resources needed to support sophisticated, high-volume online catalogs. The discussion following has been edited to reduce redundancy and irrelevancies, as have all discussions in this volume.

James F. Corey, Director, Office of Library Systems at the University of Missouri, discusses online catalog search options. Corey's presentation is divided into three distinct sections separated by lively discussion.

Charles R. Hildreth, Research Scientist at the OCLC Online Computer Library Center and author of the classic Online Public Access Catalogs: The User Interface, discusses the role of user feedback in the design of online catalogs. Much of Hildreth's presentation is based on the extensive investigations undertaken as part of the CLR-sponsored OPAC studies in 1981-83.

Joseph R. Matthews, a leading American library consultant and also a participant in the CLR OPAC studies, focuses on the importance of display formats to user-friendly systems. Matthews offers valuable practical advice to prospective buyers or designers of online catalog systems.

Michael Monahan, Sales Manager of the Library Systems Division of GEAC Computers of Ontario, Canada, and a noted online catalog designer, discusses the ways online catalog users interact with catalogs. Monahan discusses such variations as menu systems, touch terminals, natural language queries, and more sophisticated and powerful command language approaches.

Clay Burrows, President of Biblio-Techniques of Olympia, Washington, initiates an extensive discussion of online user prompts and aids with a brief yet provocative challenge paper. Burrows' comments on standards and focus on

the evolving nature of online catalog models state concisely two major themes developed during the course of the seminar.

Ray DeBuse, Manager, Development and Library Services at the Washington Library Network (WLN) and a participant in the CLR-sponsored WLN/RLG/LC Linked Systems Project (LSP), discusses the importance of linkages between online catalog systems. Inevitably, the question of standards again loomed large in this discussion.

Edwin Brownrigg, Director of the Division of Library Automation of the nine-campus University of California system, reviews options in telecommunications for online catalogs. With the recent appearance of many new alternatives to copper wires and the recent breakup of the Bell System with its attendant confusions, telecommunications is an increasingly important cost and service focus for designers of sophisticated systems.

Finally, the editor, Director for Research and Development at Blackwell North America of Lake Oswego, Oregon, offers a closing summary of issues raised in previous papers and discussions, focusing particularly on the issue of standards, which underlay much of the three-day sessions.

It is hoped that the papers and discussions in this volume will be a valuable resource for librarians and systems designers involved in the planning and implementation of online catalogs.

II. DATA STRUCTURES AND RESOURCE CONSUMPTION IN AN ONLINE CATALOG

John R. Schroeder
Jessie J. Herr

"Things that go without saying generally need to be said."
--John Kemeney, Former President, Dartmouth College

The content of this presentation is probably familiar to most system designers. Why then do we elaborate upon the obvious? Because online catalogs have a cost that may or may not be supportable in a given circumstance. We need to have this issue in front of us as we proceed to discuss other aspects of the online catalog.

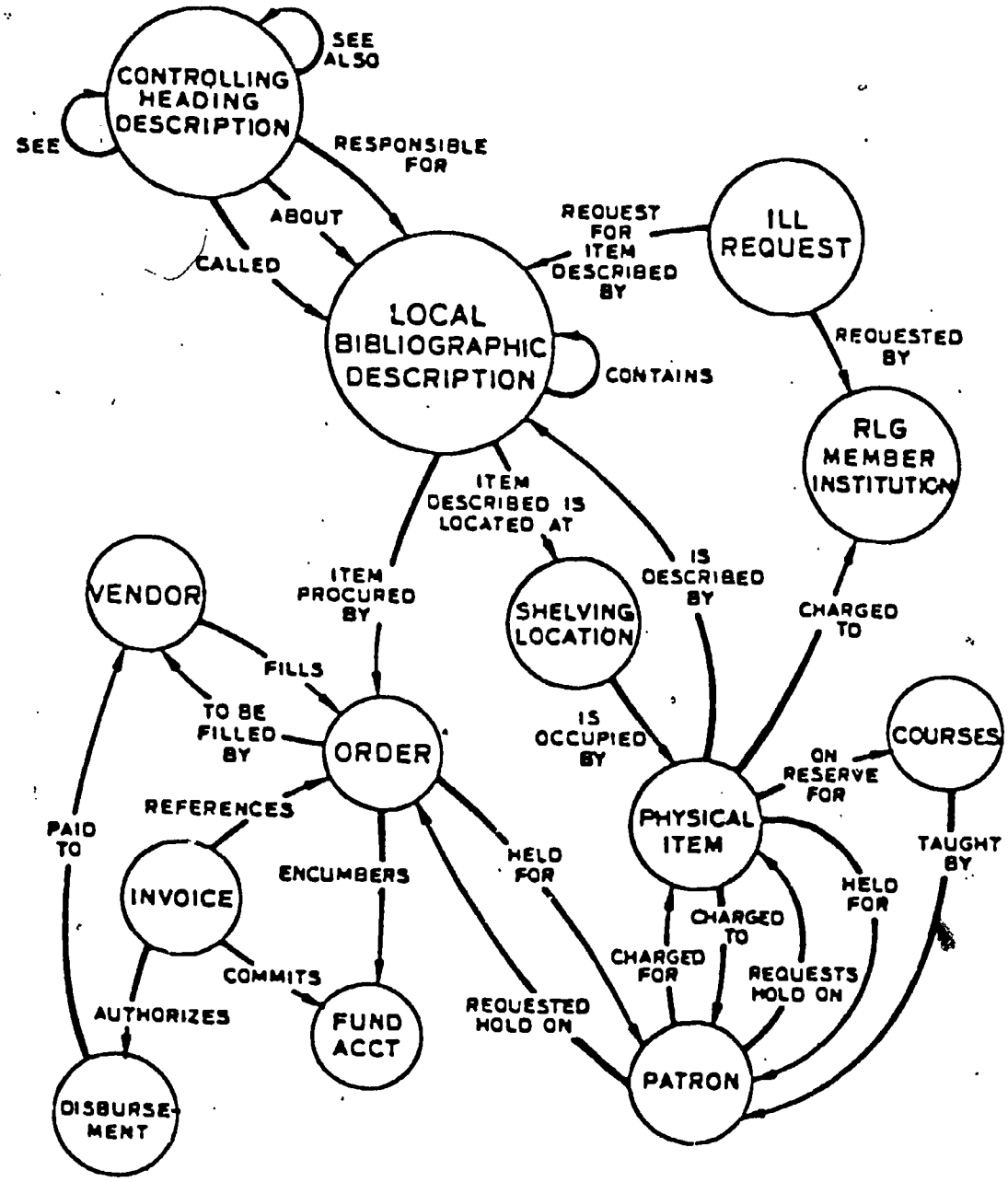
In the brief space available, we will explore what we mean by "data structures." We will focus on a particular aspect of data structures, namely, the use of inverted lists to support fast record retrieval. We will review the basic logic of Boolean search processing on inverted lists, and illustrate some scaling effects with data gathered from the Research Library Information Network (RLIN) system. Finally, we will enumerate factors that can mitigate the resource problems that may arise in a production setting.

Data Structures

The term data structures means different things to different people, so some definitions seem in order. In the formal language of database management, data structures encompass the following:

- Logical Database Structure -- The entities (items) represented in the database, the attributes (fields or subfields) of those entities, and the relationships among them. Figure 1 shows a hypothetical logical structure for a library database.
- Logical File Structure -- A subset of the above that allows for the retrieval, manipulation, and viewing of data from one or more sets of entities.
- Logical Record Structure -- The definition of the attributes of an entity, and the relationships among them.

FIGURE 1: Logical Database Structure



- Physical Record Structure -- The storage format in which a logical record is stored, together with any references to related records of the same or different type.

- Physical File Structure -- The storage format in which sets of physical records, usually of a given type, are stored to allow each to be retrieved by giving a unique record identifier or by giving one or more attribute values from within the record. This encompasses blocking and deblocking of physical records, maintenance of indexes, structures to support space allocation and deallocation within the file, etc.

- Physical Database Structure -- The totality of physical structures described above which are necessary to represent a given logical database structure within a computer storage device.

It is not our intent to dwell on logical structures. While there may be differences of opinion with regard to the abstract representation of real-world entities within a library and their interrelationships, these differences are of relatively minor importance for our discussion. Neither will we discuss physical record structures, since they are largely dependent upon taste and local design constraints.

It seems to us that the difficult data structure issue that attends the design and implementation of an online catalog has to do with physical file structures, particularly indexing and retrieval of records in the file, and the system resources that must be available to perform those tasks. This is especially true given a desire on the part of system users for the capability to specify search criteria as a logical expression of words joined by Boolean operators. (A recent survey of RLG member libraries indicated that 77% of those responding felt that Boolean Searching was a requirement. The remainder, save one, felt that this feature was at least desirable.)

In particular, discussion will center around inverted list indexing and the difficulties that arise in certain instances of its use.[1]

Inverted List Structures

Inverted list indexing is the collocation on a mass storage device of references to those records containing a given character string within the value of a given attribute. The set of collocated references corresponding to a given character string can generally be retrieved quickly given the character string. There are two basic alternatives for the structure of an index record, as shown in Figure 2.

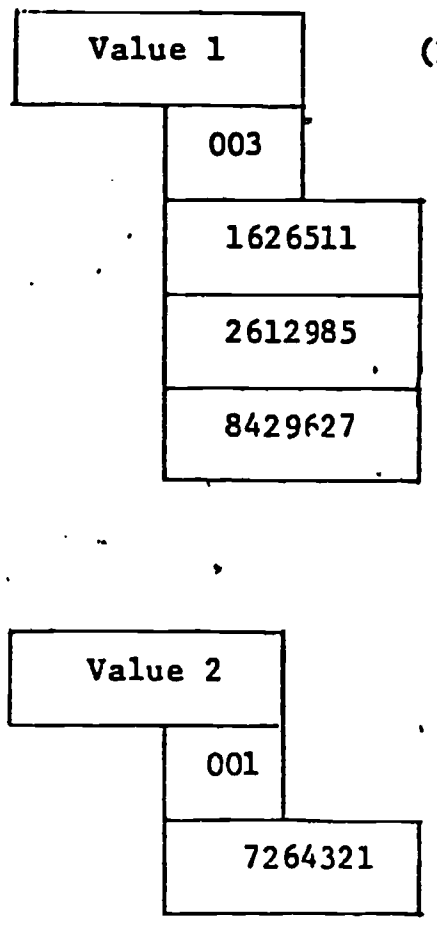
In the first alternative, each physical index record consists of a value reference pair, with the value as the record key. Since more than one record can contain a given string, the record management software must

FIGURE 2: Physical Index Record Structures

ALTERNATIVE 1

(key)		record reference
Value 1	01	1626511
Value 1	02	2612985
Value 1	03	8429627

ALTERNATIVE 2



generate tiebreakers to keep each value unique. This structure has the virtue of being supportable by a record manager that cannot handle variable length data. Some record managers that support this structure perform compression on the key values to minimize storage used for redundant values.

The second alternative is to store the value once as the key of a single index record, and to treat the references as a single variable occurrence element. This does, of course, produce a variable-length record, which not all record managers can handle.

For either alternative, the system must perform the following in order to index a given record:

1. Identify the data record attributes contributing values to the index, i.e., extract the fields or subfields to be indexed.

2. Perform required transformations on the values to be indexed. The following example is typical for a word index:

   --Convert all letters to upper case.
   --Translate special characters to null, i.e., eliminate them.
   --Translate multiple blanks to single blanks.
   --Break the text string into multiple values on blanks, i.e., extract the words.

3. Insert each resulting value, with accompanying reference to the containing record, into the index such that it is collocated with references to other records containing the same value, i.e., file the keys and references.

Boolean Operations on Inverted Lists

Given a request to retrieve all records having "John Updike" in the main entry attribute and the word "Rabbit" somewhere in the title attribute, the system must retrieve the inverted list index entries that correspond to the value "John Updike" and to the value "Rabbit," sequence each list if not stored in sequence, and perform an entry-by-entry match on both lists. Since this example is an "AND" operation, entries from either list not in the other are discarded.

"OR" operations involve similar processing, but the match is done in order to non-duplicatively merge the two lists. "AND NOT" operations are again similar, but elements of the first list are retained except for those that also exist in the second list.

Inverted Lists -- A Conundrum

Having suffered through a tiresome review of database management concepts, you will be further disappointed when we state the following: the computing resources necessary to perform Boolean operations on inverted lists are in direct proportion to the size of the inverted lists being processed. Furthermore, the average size of the inverted lists is in direct proportion to the number of records in the file.

Yet further, the more the user specifies about what he or she wants, the harder the system is going to have to work to find it. This is both backwards and unfair!

To illustrate the various aspects of this problem, we will next examine some statistics from a large online catalog.

Illustrative Statistics from a Large Online Catalog

The following information has been gathered over the past year by the RLG Central Staff while monitoring the use of the RLIN database.

1. Distribution of Record References over Values Being Indexed

There are just so many words in the Roman alphabet world, yet the size of a library's database keeps increasing. A large number of words used to describe a small number of records results in relatively small computer resource consumption. A small number of words used to describe a large number of records results in relatively large computer resource consumption.

The following statistics indicate that given an even distribution of word use across the database, inverted list sizes are quite modest.

Total number of records indexed:	4,292,474
Total number of title words:	1,679,620
Number of records per title word:	20.9
Total number of subject words:	296,551
Number of records per subject word:	91.6

A set of use counts for some non-trivial title words, which one could reasonably expect to see in search requests, begins to reveal a problem:

Ancient	10,903
Data	17,868
Development	61,375
Management	33,714
New	106,651
Social	52,440

Finally, we can examine use count distributions for title words and subject words (see Figures 3 and 4) and note that a significant number of words have pathologically large use counts. It seems reasonable to suspect that words that are popular with authors for titles are also popular with catalog searchers. We also see that library subject words provide an even more difficult picture, because of the restricted word vocabulary inherent in the Library of Congress Subject Headings (LCSH).

2. Central Processor Usage as a Function of Database Size

RLG recently increased the size of its database by 30%. It is instructive to examine central processor usage on RLG's IBM 3081 5 MIPS (million instructions per second) processor for title word searches before the addition of records and after. Searches were sampled randomly over two periods each exceeding two weeks. More than 3,000 searches were examined in all:

Average title word search before the database increase:

.76 CPU seconds

Average title word search after the database increase:

1.12 CPU seconds

It is seen that the percentage increase in CPU consumption approximates that for database size. It is further seen that the typical RLG institution running the current RLG database as a local catalog would be in difficulty, given a searching load that approximates typical card catalog usage on a 5 MIPS processor:

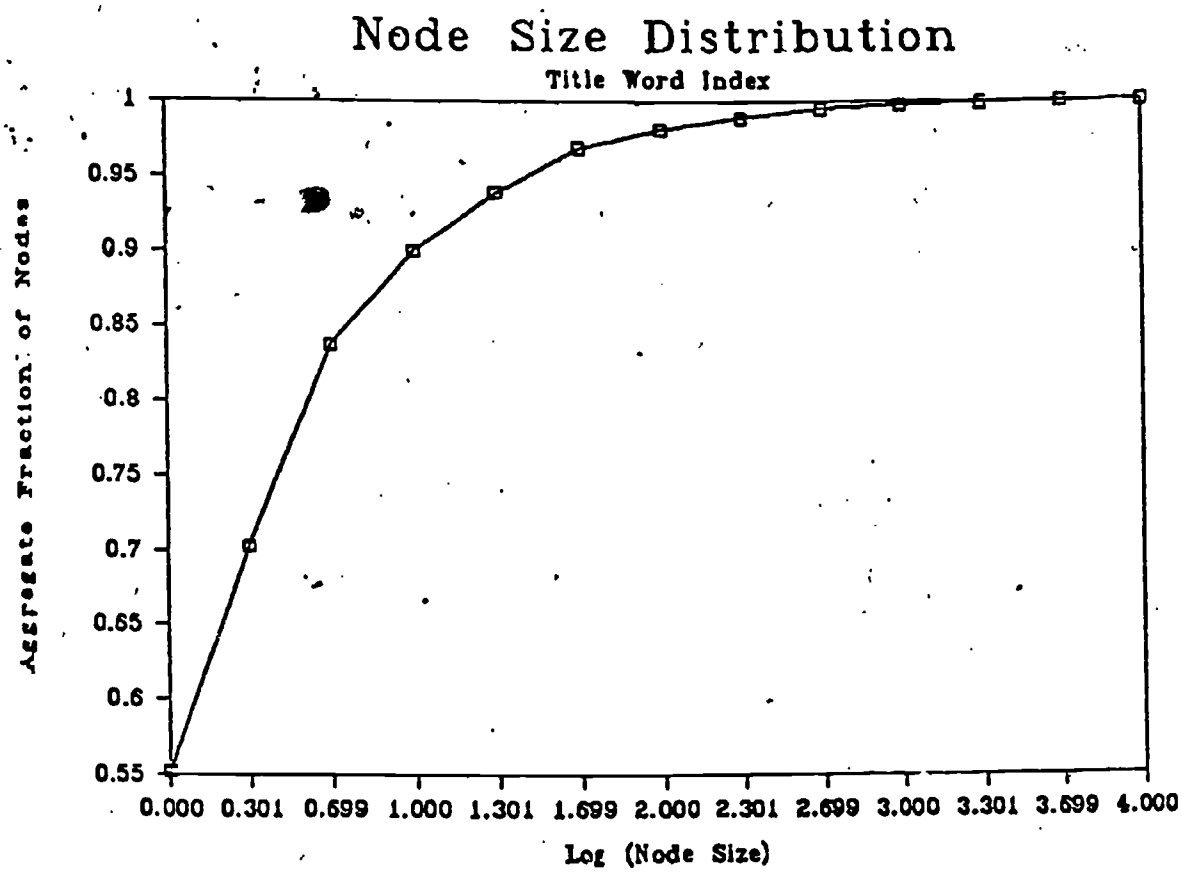
Average number of catalog searches per day:	8700
Average number of searches per hour (12 hour day):	725
Peak number of searches per hour:	3625
CPU seconds consumed per peak hour:	4060
Probable number of CPU seconds per deliverable hour:	1600

The peak/average ratio of 5 came from a study performed in 1982 by Peat, Marwick, and Mitchell at New York University. The number of CPU seconds deliverable per hour is assumed to be 60% of the total seconds; that number is further conditioned by a 35% overhead factor for the operating system.

Our hypothetical situation, while it dramatically points out the possibility of resource overcommitment, is not immediately transferable for the following two reasons:

1. The present RLG database is significantly larger than the machine-readable database in any member institution by a factor of three

FIGURE 3



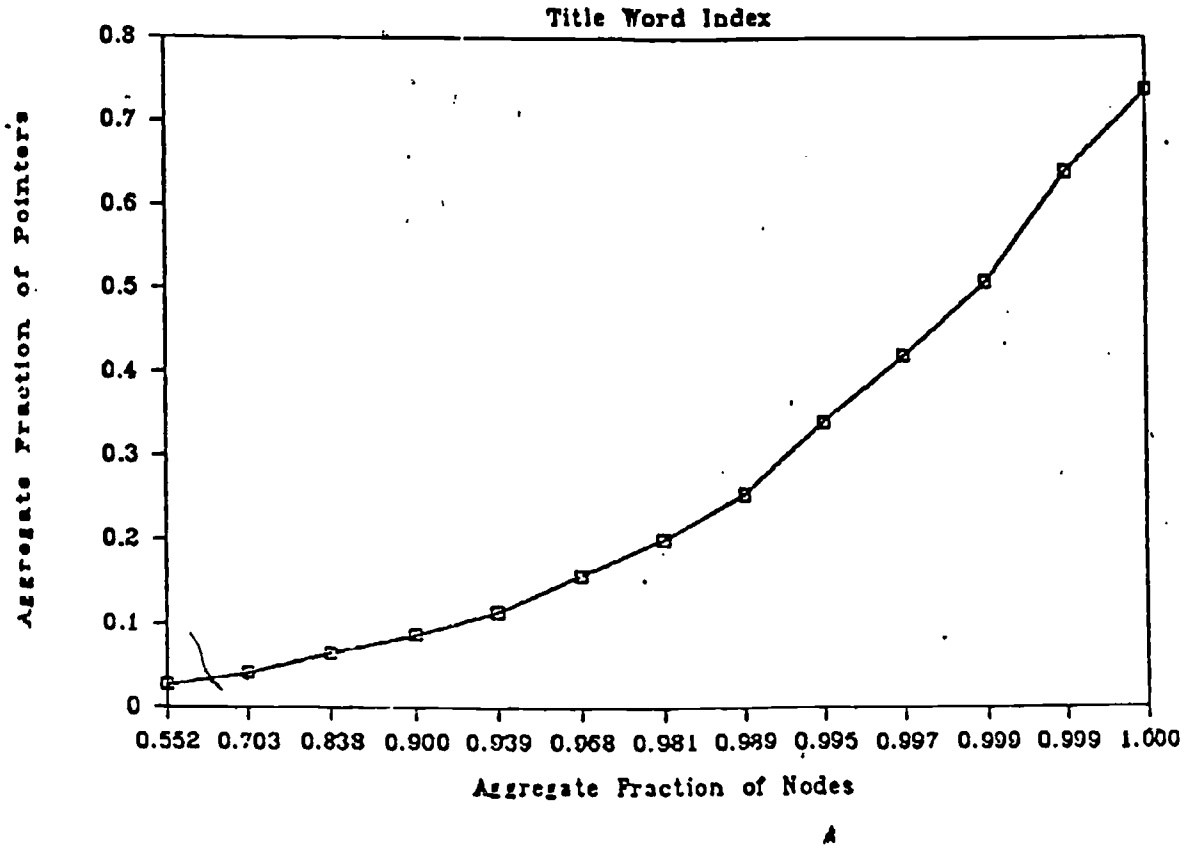
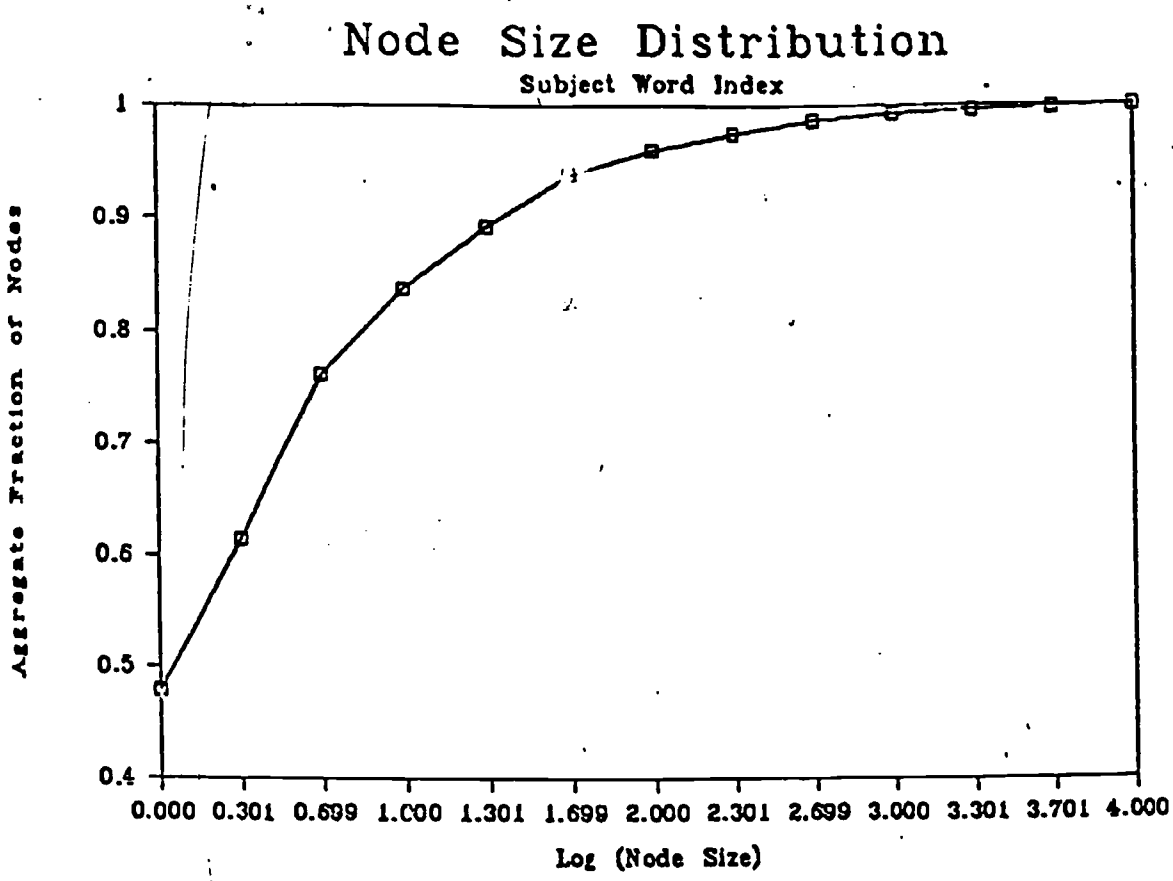
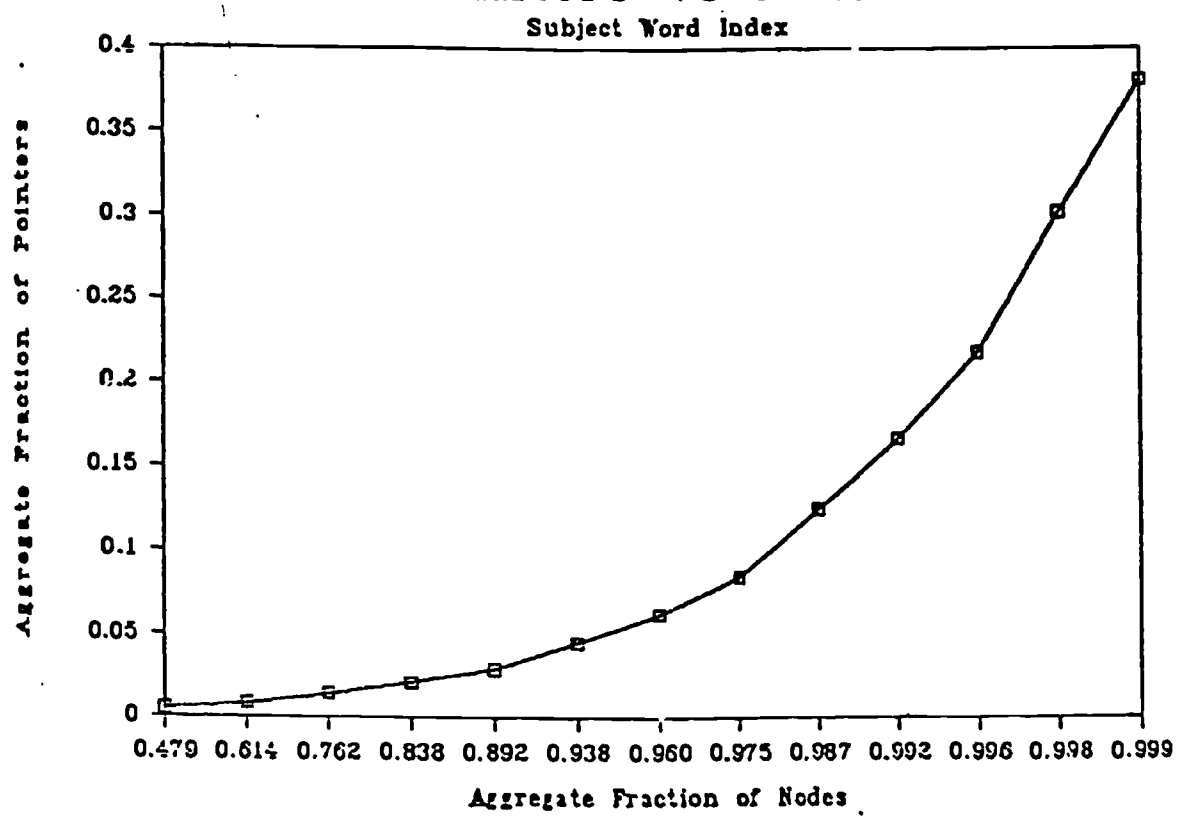


FIGURE 4





to ten. As the size of the database is decreased, there is a concomitant decline in resource consumption.

2. The present RLG software is optimized for mass updating (18,000 titles per night) as well as for retrieval. As a result, an inverted list within the RLIN database is not maintained in sequence across the entire list, but rather within a physical segment of the list. This results in implicit ORing of segments whenever a multi-segment list is processed.

   Local catalogs, which would need to accommodate many fewer updates per unit time, can be optimized exclusively for retrieval.

Staying Ahead of the Wolf

By now it should be apparent that the resource consumption associated with processing inverted lists to satisfy complex search requests against large databases is cause for worry, from the standpoint of both availability and cost. Worry need not be despair, however, for the following reasons:

1. Historically, we have seen computer price/performance improve at a compounding rate of 20% per annum.[2] Industry analysts project an acceleration of this rate, and some estimate that by 1985 mainframe computers will be priced at about $150,000 per MIPS.[3]

2. Ultra-reliable uniprocessor computers with the channel architecture necessary for the data rates required in a large online catalog are being offered now with speeds ranging from 2-14 MIPS.

3. In RLG institutions, the mean number of titles in machine-readable form is less than 500,000. Interest in retrospective conversion is building, but most institutions will not exceed 2,000,000 titles over the next eight years; one could possibly exceed 3,000,000 in that time.

The actions that can be taken to assure that demand for resources can be met are described briefly below:

1. Choose a system architecture that allows capacity upgrades.

2. Choose a financing plan that is realistic in regard to equipment amortization and replacement. An online catalog is a circulation system, where demand for resources is typically relatively static.

3. Think seriously about how the service is to be priced, if at all.

4. Perform capacity vs. demand studies carefully.

5. Choose a system that has a good user interface--poor ones exacerbate heavy resource consumption.

6. Make sure there is a way to detect valid but resource-intensive searches before they are processed, so they can be appropriately downgraded in processing priority.

7. Make sure your system has sufficient monitoring capability to allow you to predict brick walls ahead of time, to detect poor usage patterns, and to understand where the system needs improvement.

References

1. An interesting discussion of Boolean search formulations that cause slow response may be found in: Walt Crawford, "Long Searches, Slow Response: Recent Experience on RLIN," Information Technology and Libraries 2 (June 1983): 176-82.

2. J. R. Schroeder, et al., Processing and Data Distribution within the Research Libraries Information Network (Part V) (Stanford, Calif.: Research Libraries Group, 1983), 16.

3. Ulric Weil, Information Systems in the 80's (Englewood Cliffs, N.J.: Prentice-Hall, 1982), 235, 258.

QUESTIONS AND DISCUSSION

Question: John, you talked about the distribution of title words, subjects, and things like that, and obviously there are problems there, potential problems. What about indexing access points that are even more general, like language, date of publication, material type: things that qualify 60% of the database rather than 10%?

Schroeder: Well, those kinds of things are not candidates for pure inverted indexing. Essentially, you've got a couple of choices. You can tack on qualifier values to each reference in the index record, and as you process the list you can "sift" out the items whose qualifiers don't match. The problem, though, is that it inflates the size of the list. If you've got four byte pointers and 120,000 bytes worth of list and you add another four bytes of qualification information, you've doubled your I/O rate for searching.

On the other hand, if you retrieve the records themselves and check the qualification information in each, then you've got two problems: one is that you can't give an accurate count of hits before the fact, and second that's a lot of I/O operations too. I've tried both, and I think it's a matter of picking your poison.

Question: Has RLG thought about adding a second processor to its current mainframe?

Schroeder: Yes, and alternatively to get a larger first processor.

Question: What are the alternatives to inverted list indexing?

Schroeder: Someone submitted that in their written questions, too. I don't think there are any. There have been devices that have briefly appeared on the market that do associative searching, but as database sizes and search volumes grow, that technology cannot support the requisite size and speed required in the storage device. A more likely thought is a device analogous to a Fast Fourier Transform processor, which does a special thing very quickly which would execute very slowly on a general-purpose computer. In our case, we would want special hardware to perform Boolean operations on lists.

I think we need to retain our perspective, though. 100,000 pointers sounds pretty scary, but I can remember how scared we were in 1972 when we built a 50,000 record database. Today, that is minuscule, and I have a feeling that we're going to look back at 100,000 pointers ten years from now and say, "Gosh, that wasn't so bad."

Question: I want to know, given the situation RLG has run into, whether your experience points to living with slow response time, or to decreasing the number of terminals.

Schroeder: I think the latter is preferable. People would far rather wait to get on and have some decent service when they get there. We have, of course, been in that situation.

Question: I know the history, but have you learned anything from it?

Schroeder: He's got a nasty streak in him.

Question: So one thing to be added to your list of recommendations is the ability to vary communication ports in and out of service?

Schroeder: Well, some of the communication switching gear now has queueing capability. People can "line up" behind the busy ports, and the system will beep at them when their turn comes.

Question: What about local area networks, and what about their effect on resource consumption generally?

Schroeder: We found in the Carnegie survey that a significant number of research libraries are going to interface their processing capabilities to a LAN, and are talking about hundreds, if not thousands, of workstations. And the people who own those devices fully expect to be able to access the catalog, and in fact the catalog is often felt to be the most valuable resource on such a network.

Question: Do you advocate the movement of searching capabilities into these workstations, and if so, will this relieve capacity problems in the library computer?

Schroeder: Well, if you look at the resource consumption currently associated with syntax checking and conducting tutorial and "help" dialogs, you will find them insignificant in comparison with that devoted to processing the inverted lists. But in the future, the application of "expert systems" techniques, natural language front ends, and the like may change the proportion of resources devoted to the interface into something much larger.

Question: What about the use of large multiprocessing arrays of small computers?

Schroeder: You are processing many lists for many users, but in parallel. I'm suspicious of this kind of complexity, though. I guess I am a conservative, and it doesn't strike me that the single processor approach ought to be discarded just because it's unimaginative. Single processors exist today that will do the job. For the present they are unaffordable, but within three to five years, we will see what we need at a price we can pay. The software to handle large database applications on these complex multiprocessor systems is not in existence.

Question: How long does it take to download an RLIN terminal? I heard a story about somebody who turned off a terminal at Davis, and when they turned it back on, it took hours to get back on the air.

Schroeder: It usually takes about three minutes. Maybe they had a dirty line. If there were a lot of retransmission of error frames going on, it could take that long.

Question: Do you see any potential for segmenting the database to give the user only that part that interests her or him?

Schroeder: I suppose that would be possible, but I don't know how to do that.

Question: In the context of really long inverted lists, how do you deal with the problem of truncation and partial matching?

Schroeder: Truncations are handled by implicit ORing, which is the only way I know to handle them. We go into the index tree at the point where the root first occurs. The index is then read sequentially, and each record having that root is ORed into the result. Floating partial match strings may be applied to each index key prior to the ORing.

Question: What is the optimal ratio of index and data space? When do things start to run into trouble?

Schroeder: I can't answer that in general, since all systems are different. I don't think that ratio is critical, except as it relates to the cost of storage.

Question: Walt Crawford did an elaborate set of analyses of types of searches, and I wonder if you have made any changes as a result of his findings?

Schroeder: Well, I think we've tightened up our liaison with users a little bit, because our users, remember, are library staff at the moment. A lot of problem searches turned out to be accidents. We happen to have two index mnemonics: title phrase (TP) and title word (TW). Guess what people were doing?

Furthermore, we don't have long search detection a priori, because we don't maintain posting counts up front in the index record. We've done a projection which shows that in 1992 we'll have 32 million records, so some reworking of the access structure to provide this capability among others seems called for.

Question: You talked about how you made the system less efficient for retrieval in order to have updateability. Is it possible to take a static

segment of the database and optimize the hell out of it for retrieval, and then have another more volatile portion optimized differently?

Schroeder: I don't think that is necessary, given the number of updates that will be generated in the typical institution.

Comment: I want to point something out for those who are enamored of the concept of storage getting cheaper. What isn't true is that the devices take any less space and generate any less BTUs.

Schroeder: I'll take mild exception to that. 3380s generate a third the BTUs that 3350s do, and the footprint is less than half.

Comment: If you look at figures three and four, the curves are essentially the same shape. What you said, John, was something to the effect that those big ones are the ones that people are going to want to search most often. What we have to do, I think, is to really get some specificity about which words and word combinations are high access and which aren't by empirical data gathering. These should then be treated separately.

Schroeder: That's right. There are four or five large databases that would be a perfect environment for that kind of research.

Comment: John kind of passed off the need for real-time update. I would agree with that if we're talking about patron access. But I know some acquisition librarians who would strongly disagree, because of the possibility of ordering duplicate copies.

Comment: There are some compromises that can be taken here, where real-time might not necessarily mean at the stroke of the enter key. It may happen in background, and be available in a relatively short period of time.

Schroeder: That's a nice compromise. Another is to have two separate databases: one small one for new records, and the large one. You check the small one before accessing the big one. They can be merged offline. But I think that the small amount of updating activity in a local environment makes real-time indexing more possible than in a centralized environment. You do need some crash and backout logic, but that's pretty commonplace now.

Question: Getting back to the indexing problem, wouldn't it be possible to generate both phrase indexes and word indexes and process a search against both?

Comment: Yes, how can you take advantage of what seems to be a fact that one-fifth to as many as half, but at least one-fifth to one-third of our users know the exact author's name and the exact title? How can we take advantage of that and not search title phrases as title words?

Comment: One idea is to try an exact phrase search first. If hits are produced, then allow the user to indicate whether a subsequent word search is necessary.

Schroeder: With a database of our size, that would cost three of four accesses.

Comment: Not a big deal.

Question: Have you people gotten crafty about deciding which word is bigger, and optimizing the search request accordingly?

Schroeder: Well, since we don't store posting counts in the index...

Comment: That would buy you the potential of avoiding a lot of comparisons.

Schroeder: Yes, well, we do optimize the search requests insofar as we can. But not putting those posting counts in has really turned out to be a dumb thing for us to have done.

Comment: I think the important thing, and this cuts across a whole spectrum of specific problem areas, is to consider the role of heuristics that can be put into the system. We can weave these heuristic techniques into the interfaces to do the kinds of things that correspond to what user behavior is like, and optimize the system resource handling accordingly.

Question: Have you looked at the possibility of using many processors, each with a "local" online catalog, but having the machines in a central location?

Schroeder: For RLG, the issues surrounding that idea are programmatic and political more than technical. When you stop and think of a research institution depending upon an online catalog as the primary means of accessing the collection, you realize that no institution would tolerate having so important a tool outside their direct control.

Comment: From this discussion, it is clear that a number of different things taken together are needed to solve this problem, rather than the approach you have suggested, more focused on a big machine.

Schroeder: One can spend a lot of money developing a sophisticated combination of solutions, and by the time you have it working, technology has rendered it obsolete.

Comment: It may be complex for us, but it may produce more simplistics for the users.

Schroeder: All I'm saying is that there is a crossover point. Complex solutions take a long time to design and to implement.

Comment: One of the terrors, I think, of this sort of processing is something you talked about...that the more the user brings to the search, the worse off he or she is, depending obviously on the search.

The problem is that the poor person doesn't know they have just walked into your twenty-word land mine. Has anyone experimented with conditioning the user by reporting back which searches are expensive, which are cheap, and why? Libraries don't have infinite resources, so I wonder if we may have to start to tell people, "Look, you're going to have to learn the peculiarities of what we do. It's not what we want to live with, but this is expensive, so don't do it." Or maybe, "Be careful when you do it."

Comment: RLG has started to do that, as I understand. They have an option where it will feed back the cost of the search to you.

Comment: Haven't we learned anything from the behaviorists about positive reinforcement? What you're talking about is punishing for bad searches. What about the other side? Reward them!

Comment: Refund the quarter. (laughter)

Schroeder: Reward the person who enters "marine ecology" instead of "journal marine ecology."

Comment: Another issue having to do with data gathering on user behavior is confidentiality. Those of us who administer system use are provided with examples of dangerous and resource-consumptive searches, but we have no way of telling who conducted them.

Schroeder: We took this up during the BALLOTS days with the University Committee on Privacy at Stanford. They gave us two guidelines, as I remember: ask the user's permission to log the session, and don't use the data for behavioral research that can be related to any individual or group of individuals. In other words, only use the data to improve the system. What you can do is ask the user his or her permission to log the session. You're home free if the user says yes, which most of them do--particularly if that's the default. (laughter) We've gotten a lot of insight into what people do right and wrong from those logs.

Comment: In operating systems, one has the concept of a "working set," which is a set of memory pages empirically determined to be popular. At one point in this discussion, we started talking about analyzing our database. That really isn't the point. We should be analyzing our searches, and empirically determining what is popular in the database--the working set for the database.

I have this conviction that people who understand system architectures of basic operating systems have been down this path. Basic operating systems have learned to manage paging tasks and things that are beyond what I think our applications levels are doing today. They haven't quite studied our systems yet. The people who did MVS/XA have something to offer us, if we can get them to study the problem.

Comment: One of the other problems, I think, is the fact that we have created a difficulty before the thing even gets installed, because we have so raised expectations. People expect somehow that the online catalog is going to be a marvelous tool capable of delivering everything under the sun. They have no idea what the costs are. Don't you think some advanced preparation is necessary?

Schroeder: Absolutely. It's like climbing on a train for free and having to pay to stay on. Once people install and demand builds, there is no way to back out gracefully. One can only meet the demand. People had better understand that beforehand: what the demand curve is likely to be and what kinds of support dollars are necessary.

Comment: And yet everyone, when they're buying the system, asks for everything. Very seldom do we let them know that there is a real cost involved.

Schroeder: I happen to think that even though what we are talking about is costly, people are going to be willing to pay for it. But they have to know up front. I have talked to several groups of chief financial officers in RLG institutions about this. They don't back away, but they want to know what kind of long-term commitment they're signing up for.

Question: What kinds of symptoms of resource overcommitment should we look for early on? Where have the symptoms cropped up first in your experience? Is it in the processor? Is it in the channels? Is it in the disks?

Schroeder: I can't answer that question in general. We've got excess channel capacity, and we've pushed the CPU a bit, but our principal problem has been memory. Someone else will experience channel problems first. It depends on the system and the load.

Comment: There should be standard benchmarks that could be run so people could adequately judge how a system is performing. Not a standard as such, but at least a baseline against which judgments can be made.

Comment: Terminal simulators. That would be one of the good things to put in any specification.

Schroeder: An amazing number of systems do not have those things. And with those that do, it takes 85 man-years to write the script. Some time back, we took our command log and reduced it to a graph of state transitions, with probabilities associated with each edge. Since we also added statistical think time, we really had a semi-Markov process. We used this to drive the system to a level of several hundred terminals.

Question: Did it help?

Schroeder: We learned a lot. Problems associated with response all have to do with serial use of a resource. You discover one bottleneck and fix it only to find the next one at a later time. Having such a tool allows you to predict where the next one will be, and when.

Question: John, I read an article recently which indicated that the IBM 308Is of today are the micros of tomorrow. We are making incredible strides in the technology.

I'm an advocate of distributing the catalog across several processors. That presents some interesting challenges for data structures, especially where a union catalog is involved. How important can a network library feel the union catalog is, as opposed to being able to send searches to the other local catalogs in the network?

Schroeder: Take RLG as an example. We depend upon a central database to support a resource-sharing program. The database potentially contains location information for all material in the consortium, and locations are needed to support interlibrary loan.

Now, we could physically replicate this union database, and ship weekly copies to the participants. That's an approach I think we'll be taking within the next five to ten years. Meanwhile, we have to choose between having a central database, or alternatively broadcasting searches, which means that hypothetically every machine in the network has to be able to handle the combined search load. I really don't think this second alternative is viable, unless searches can be routed selectively and intelligently using a conspectus of collection strengths.

As far as micros versus mainframes is concerned, I hold no brief for one or the other. In any case, you've got to provide a lot of computing power and I/O bandwidth to process searches on a large database. How one does this is a matter of taste, and of the constraints one operates under.

Comment: Another attractive online catalog feature is sorting. The user of the catalog can go on to want their data presented on the screen sequenced by main entry, title, I don't know what? When it comes to that, are there any indexing tricks that reduce the need for sorting?

Schroeder: I don't know of any that are general. I was afraid someone would ask that.

Comment: Some studies seem to indicate that users don't care about sorting a great deal of the time. But there is the guy who's going to want to print off a big bibliography and he's going to want it by subject and who knows what else. So we will negotiate with him, through the interface. We tell him it's going to take a while, and we ship it to an offline process.

Comment: It seems to me, again, that every one of the things we talked about are old problems to somebody else. It would benefit in general to look at the online retrieval systems, because they have tackled and solved a lot of these things that appear to us as new problems.

Question: John, building on that and the questions that were raised earlier, do you dismiss the idea of intuitively obvious subsets?

Schroeder: No, I didn't mean to do that. I may sound like somebody who's just waiting for more crunching capability. But I do think that something like this needs to be tried to hedge the bet since the demand may outstrip our ability to brute-force it.

Question: I'd be interested if anyone has comments about a time-segmented catalog, where some user initiative is required to move to the next older segment. Perhaps another quarter in the slot, to go back three more years.

Schroeder: I like the descriptive phrase "working set" applied to the often-sought portion of the database that we are trying to isolate.

Comment: The idea of a segmented catalog is a difficult one to act on. Because as soon as you think you have a decision worth making about how to segment, someone on the other side of campus objects.

Schroeder: And what would work at Columbia would never work at Stanford, or at Columbia either in the next instant.

III. SEARCH RETRIEVAL OPTIONS

James F. Corey

Today's online library catalogs differ from one another in many respects. These catalogs exhibit a wide variety of dialogue techniques between person and machine. They exhibit a variety of display formats for bibliographic information, and the underlying file structures of the catalogs are different. Online catalogs also show a considerable variation in search retrieval options that are available to the online user.

This paper enumerates search retrieval options that online catalogs currently have, and offers questions about the relative importance of various search options or techniques for implementation of some search options. This paper assumes that there is a way, a form of dialogue, for the user to express his or her search to the computer and that the computer responses will be in some format that the user can understand. This paper will not treat the topics of dialogue, file structures, or display format and content because they are being covered by other papers at this conference.

Search retrieval options are intimately connected to the file structures of an online catalog. The term "file" is used here in its broadest sense, meaning a collection of related records. No physical structure is implied. There is usually more than one file structure that can be used to achieve a given search feature, but in one form or another, online catalogs must have the following files:

1. Authority Control File. An authority control file is necessary to assist in controlling names, series, and subject headings. This file should be compatible with the MARC authority formats and should contain cross-references (SEE, SEE ALSO, SEE FROM, SEE ALSO FROM), scope notes, and verification notes. For subjects, the authority file should accommodate and distinguish multiple sources of headings, i.e., LC, NLM, Sears, and locally established controlled vocabularies.

2. Bibliographic File. The bibliographic file should contain MARC compatible bibliographic records for all types of library materials.

3. Holdings File. The holdings records should have enough detail to distinguish each physical item. The system should have the ability to store and display a variety of holdings statements, from the briefest summary to the detailed item level, for any format of material.

The authority, bibliographic, and holdings files, if all are indeed physically separate, should be linked together in some way so that search requests can proceed from authority to bibliographic to holdings records and vice versa.

In order to provide a framework for a discussion of search options, it might be useful to distinguish two broad types of searches that the computer may perform. One type of search is the "scan search," in which the user is presented with sorted lists of like bibliographic data. Common examples of scan searches are searches by author, title, subject heading, and call number. These common fields are usually organized in a quasi-alphabetical order, using mostly word by word filing but having various exception rules depending on the field and the online catalog system. The scan search is the closest search in the online catalog to searching that is performed in card and book catalogs.

Scan searching is sometimes called browsing, but the word "browse" is also frequently used in the literature to mean a more complex type of searching in which the user retrieves some records, scans through them, then jumps to other records and inspects some of the latter and then goes to still other records.[1-4] Frequently the browsing user will use information from records previously retrieved to spur the search for successive records. In this more complex sense of browsing, scan searching is only one component of browsing. The use of the word "scanning" in this paper to mean looking through alphabetically related records is consistent with the use of the word in an OCLC research report by Markey.[5]

The second general type of search is the "set retrieval" search. In a set retrieval search, the user provides search data along with information on how the data are to be used, and the system looks for the set of all records that meet the search criteria. Examples of set retrieval searches are a search for all authors with the surname "Williams," a request to retrieve all books by a given author, a search for all books in a monographic series, and a request to display all copies of a book at a branch library.

The set retrieval type of search can be further subdivided into two subordinate types of searches--the keyword (or free text) search and the phrase match (controlled vocabulary or heading) search. While in most keyword searches the order of the supplied words is unimportant, a controlled vocabulary search requires the user to supply an exact sequence of terms, usually beginning with the leftmost word in the heading. For the controlled vocabulary search, the user must be aware that the order of terms has to be correct. For the keyword search, the user does not have to know that the order of terms is unimportant in order for the user to initiate a search, but the user will probably be surprised by the results of the search when items are retrieved in which the searched terms are not adjacent or in the order expected.

There are several observations that can be made about scan and set retrieval searches. First, a scan search requires that the search terms be entered in correct order beginning with the first word of the search argument. Keywords cannot be used in a scan search. Second, Boolean logic is not used in a browse search, but is limited to set retrieval searches.

Third, browse and set retrieval searches can be, and usually are, intermixed during a user's session. The user may scan through a list of authors, see a relevant author name, and ask for all books by that author. Or a user may ask for (the set of) all subject headings that contain a given keyword, pick one of the headings, and begin scanning subject headings beginning with the selected heading.

Fourth, indexes that contain only one term may or may not be scannable, but they are more likely to be used for set retrieval. For example, the ISBN, ISSN, and LC card number indexes contain "one word" values and have low utility for scanning. They are more commonly used for retrieving all records having the supplied number.

Fifth, a scan search is, in the strictest sense, a set retrieval, but the set is composed of approximately one CRT screen's worth of records that are alphabetically adjacent with respect to one field within the record. For this paper and discussion, set retrieval is not meant to include alphabetically adjacent records. This point is important to make when one considers the similarities of some possible screen displays. Figure 1A shows the hypothetical result of a scan of an author file using the search value "Ryan." Figure 1B shows the result of a set retrieval search for all authors named "Ryan" in the same author file. There is essentially no difference between the two responses. For less common names, the difference between the two types of searches is more apparent. Figures 2A and 2B are displays for a browse and a set retrieval search using the search value "Rick."

In the following sections, various search options for scanning and set retrieval are described. The phrase "the online catalog should have..." is used frequently to introduce one search option after another. The use of the prescriptive, "should have," is not meant to indicate that the value of a feature has been solidly demonstrated by online catalog use studies. On the contrary, online catalog studies to date have not focused on the detailed analysis of the effectiveness of one search technique compared to another. The notion that online catalogs "should have" various features is, at this early state of online catalog development, intended to mean that the features seem to be, to one degree or another, important to the successful operation of online catalogs, are in fact already available in one or more online catalogs, and will undoubtedly continue to be used in online catalogs until better search techniques are brought to light. It may be a truism, but until we know of better search methods, online catalogs should have the features that librarians currently think are best, and these are the features that are in fact being implemented in online catalogs.

FIGURE 1A: Browse on common name

browse on author: ryan

1. Ryan, A. H.
2. Ryan, A. J.,
3. Ryan, A. N.,
4. Ryan, Abram Joseph, 1836-1886.
5. Ryan, Adrian.
6. Ryan, Agnes,
7. Ryan, Alan.
8. Ryan, Alfred Patrick.
9. Ryan, Allan J.
10. Ryan, Allen J.
11. Ryan, Alvan Sherman, 1912-
12. Ryan, Andrew Field.
13. Ryan, Anna I.,
14. Ryan, Anne.
15. Ryan, Anne, 1889-1954.
16. Ryan, Anthony S.
17. Ryan, Anthony S. Formative evaluation.
18. Ryan, Arnold W.

FIGURE 1B: Set Retrieval on common name

search for author: ryan

found: 734 records; first 18 displayed

1. Ryan, A. H.
2. Ryan, A. J.,
3. Ryan, A. N.,
4. Ryan, Abram Joseph, 1836-1886.
5. Ryan, Adrian.
6. Ryan, Agnes,
7. Ryan, Alan.
8. Ryan, Alfred Patrick.
9. Ryan, Allan J.
10. Ryan, Allen J.
11. Ryan, Alvan Sherman, 1912-
12. Ryan, Andrew Field.
13. Ryan, Anna I.,
14. Ryan, Anne.
15. Ryan, Anne, 1889-1954.
16. Ryan, Anthony S.
17. Ryan, Anthony S. Formative evaluation.
18. Ryan, Arnold W.

FIGURE 2A: Browse on less common name

```
browse on author:  rick

    1.  Rick, Alan,
    2.  Rick, Forest O.
    3.  Rick, Jay.
    4.  Rick, John H.
    5.  Rick, John W.
    6.  Rick, Lilian L.
    7.  Rick, Wirnt.
    8.  Rick Friedberg Productions.
    9.  Rick Trow Productions.
   10.  Rickabaugh, Carey G.
   11.  Rickaby, Joseph John, 1845-1932.
   12.  Rickard, Andy.
   13.  Rickard, Anthony Robert.
   14.  Rickard, Bob.
   15.  Rickard, Clinton, 1882-1971.
   16.  Rickard, D. T.
   17.  Rickard, David T.
   18.  Rickard, David T. (David Terence), 1943-
```

FIGURE 2B: Set Retrieval on less common name

```
search for author:  rick

found:  9 records; all displayed

    1.  Rick, Alan,
    2.  Rick, Forest O.
    3.  Rick, Jay.
    4.  Rick, John H.
    5.  Rick, John W.
    6.  Rick, Lilian L.
    7.  Rick, Wirnt.
    8.  Rick Friedberg Productions.
    9.  Rick Trow Productions.
```

Scanning

The online catalog should allow the user to scan authority records, bibliographic records and holdings, or indexes to the three files. The common fields in these files that should be amenable to scanning are author, title, series, subject headings, and call numbers. The user should be able to scan both forward and backward. To start a scan, the user should have only to specify the desired type of field and key as many characters or words as the user thinks are needed. The search argument should be entered in correct order beginning with the first word because scan is a phrase match type of search. No form of explicit truncation should be needed; truncation should be automatic.

When scanning author, series, or subject headings, it would be helpful if the catalog displayed the number of bibliographic records for each entry. When scanning authority records, the user should be able to request bibliographic records associated with a displayed heading with a minimum of keying, and certainly without having to type the heading of interest. When scanning bibliographic records, the user should be able to request easily the holdings for one or more works. If a heading in an alphabetical list is, in fact, a SEE reference to one or more authorized headings, the system should have the option of being able to include the authorized terms under the unauthorized term (probably indented and prefixed with the word "SEE"). If a heading in the displayed list has a SEE ALSO relationship to other headings, the user should be made aware that the SEE ALSOs exist, but the headings should probably not be displayed until the user requests them.

The online catalog should permit multiple call number indexes so that, in effect, the system can have any number of online shelflists. The system should permit shelflists for branches and locations within libraries or branches (e.g., Reference). The system should provide separate shelflists for each classification system used by a library or branch.

Before going on to set retrieval searches, it may be worthwhile to stop at this point and discuss some of the questions about scanning options.

References

1. Mark S. Fox and Andrew J. Palay, "Machine-Assisted Browsing for the Naive User," in Public Access to Library Automation, Proceedings of the 1980 Clinic on Library Applications of Data Processing, (Urbana: University of Illinois, Graduate School of Library and Information Science, 1981).

2. Charles R. Hildreth, "The Concept and Mechanics of Browsing in an Online Library Catalog," in Proceedings of the Third National Online Meeting (Medford, N.J.: Learned Information, 1982), 181-96.

3. Wilfrid F. Lancaster, Information Retrieval Systems: Characteristics, Testing, and Evaluation (New York: John Wiley & Sons, 1968), 181-182.

4. Saul Herner, "Browsing," in Encyclopedia of Library and Information Science (New York: Marcel Dekker, 1970), 408-15.

5. Karen Markey, The Process of Subject Searching in the Library Catalog, Final Report of the Subject Access Research Project (Dublin, Ohio: OCLC Online Computer Library Center, 1983).

DISCUSSION OF SCANNING OPTIONS

Corey: We will change the discussion format used by John Schroeder last night. John's discussion was something like a star network with the audience asking questions and John giving witty or wise answers and sometimes both, and then another person asking a question. For my part of the program, we will use the ring typology. Someone asks a question and we will have it answered by the person on the left of the questioner. I say this because I have read the questions you have submitted and they are very difficult to answer. Certainly, I do not have the answers.

I decided to put scanning first in the paper because it is not getting any coverage to speak of in the literature. Scanning is just somewhat taken for granted, and everybody is talking about the fancier, more elaborate searches. But I think scanning has a lot of power, and it has a lot of simplicity, and we ought to give it a few minutes of review.

Scanning bibliographic records presents a special problem because it does require that the records be sorted in some way, and usually one has done a set retrieval search prior to the sorting and before one begins to scan the records.

Scanning is not an interesting computer science issue. From an algorithmic point of view, it is fairly dull and elementary, and I don't know if it needs a lot of research. I think the question that might need some research is when the option to do scanning searches should be presented to the user.

Now that's basically all I want to say about scanning, except that there are some questions about it--and I thought we might discuss these questions for a few minutes before we go on to set retrieval. Here's a question that one of you submitted: "How important is browsing capability and, if important, how should it be implemented?" Because of this ambiguity over the word "browse," I'm not sure whether the questioner meant browsing in the more sophisticated sense that Charles Hildreth has written about, which means "hunting around," that is, finding some records and then going from those records to related records, and then going from these other records to even further records by following tracings or using other ways of finding links to records. So I don't know, given the question, that the questioner was limiting this question to scanning--but I thought he probably was. So the question is--how important is scanning?

Comment: I don't think it's that important whether we call it "browsing," "scanning," or anything else. I'd like to specify the question a little bit more. I think we all know about this kind of activity, and you described it well, whatever label we give it. I think two subquestions become very important--what is the actual activity, what is it we're browsing or scanning at any given moment with the screen as the representation of that

file and the structure and the records in it, and so forth, and why are we undertaking that particular kind of browsing or scanning?

I do prefer to keep browsing as a higher level activity or concept, if you like, so the word "scanning" doesn't bother me. But there are some important distinctions in what it is scanning, such as headings, whether they come up automatically or will we use them to expand or command or whatever else--or whether scanning is for short or long bibliographic representations. I think there are differences between those activities, and they serve different purposes. I think we need to specify the question, at least that part of it.

Comment: In your discussion, it was suggested that this technique be used for authority files. I might suggest that scanning is a technique that works best with headings in controlled situations, when we control the way authors file, when we control the way subjects come together.

Comment: I'd like to answer my first question if I can, why is it important to do what I called the first kind, the scanning of headings, if you like. Well, there are some traditional problems that it takes care of--one for the user and one for the system, to improve the efficiency and effectiveness of both.

There is the entry vocabulary problem, we're all aware of what that is. If you're committed--not even coerced by the dialogue structure, you scan the headings--authoritative versions, and important cross-references are in there, and you're helping with the problem of entry vocabulary, that's the basic answer.

But you're also--especially if you force that scan on the naive searcher, and perhaps the expert, too--you're going to start narrowing at that point, especially if it's controlled vocabulary, tree structure or whatever. You might have the concept, well, they've seen subheadings, and there are other ways it can be done, there are guides to the scanning list. So why scan? Why is it important? I think that was the question.

It helps with the entry vocabulary problem, and it encourages the searcher to narrow, refine right from the start, even if they're not familiar with the operation of the Boolean logic and set retrieval.

Comment: I have all kinds of mixed thoughts about this whole thing. DOBIS uses this entry technique for all access points, for all parts of the system, not just the bibliographic parts, but also vendors and library funds. The feeling was that that kind of technique is very close to the kind of searching that somebody does now in card catalogs. They page through until they find an entry that looks nice. This technique probably solves 80-90% of all searches, almost immediately, because they find the entry they want, there are 5 or 10 or 20 documents underneath it, and they can look through those and

that ends the search. They don't need to make this step to the Boolean and more sophisticated searching techniques.

It does cause problems, at least in one area, and that is that librarians don't believe in alphabetical order. The name "McAllister" is filed before "machinery." Which is filed before "McMann," or something like that. And these filing sequences are ones that our real live users don't really understand.

The problem gets more difficult when you get into foreign language catalogs, where there are a large number of umlauts, haceks, and other strange characters which don't file as our hex characters recommend, and in fact there are hidden codings behind many of those.

I don't know whether you make the argument, okay, we'll file in alphabetical order, period--and leave the librarian hanging there. Or whether you end up saying okay, we'll file your way, librarian, and it's up to you to somehow come to grips with this weird sequence of access points.

Comment: A corollary problem to consider might be recognizing that often, the user may want to start at a known point because the user feels that he knows for certain the beginning word or letter or sequence, what have you--but they are uncertain about the rest. If you provide scanning facilities, then you want to consider providing pattern matching facilities as part of that. Moreover, if you think of specific types of scanning, like for authors, and you know that quite often the user may have started within an erroneous spelling or they are not sure of the spelling, then you may want to consider sophisticated automatic spelling error detection and correction types of facilities built into the system.

Comment: It might be worthwhile to throw in a little word about the experience we had with DOBIS, where we started out with only scanning--that's really all we had in the beginning, and it's the basic thing that we have today.

We got--not really very early on, but not too long after we actually started using it, I think--requests for Boolean operations, but not very many. You can get along an awful lot without it. You don't really need it. We always put everything, all users' wishes down on lists, and we sort of prioritize them, and we try to get them done some time. This grows.

Finally, I took a couple of weeks some time, and just threw Boolean logic into our system. It wasn't a big job, as a matter of fact. It took a little working around with algorithms, and it's not a very fancy implementation, but it works.

Then, of course, we gave it out to our users--and the question is then, do they use it? Well--yes and no. The 8% is probably a pretty good

indication of what happens. I think probably about 90% of our users never touch it at all; and we do have users doing searches in some of our institutions. Most never touch it at all, but some of them do use it and use it well. But they don't find that that is the way to do the job. The way to do the job in our system is by the scan search. Then if you come up with a long list, then you limit it, then you cut it down. But not until that point. You don't start from there, at least in our experience.

Corey: If you are to have a scan search, there are some questions in your implementation you have to have answered, and I have a list of them here, and I'll just read them. If one of them strikes your fancy, interrupt. I would like to at least cover the kinds of questions because I think there are issues that we have to answer, at least individually, in our systems. And we have to back our answer with the dollars we spend on development to achieve the function.

Here are some of the additional questions on scanning. What is the nature of the indexes that will be browsed? Specifically--(a), will the indexes contain full headings without regard to length, or will the indexes be truncated headings of, say, a length that fits conveniently on one line of an 80-character CRT? Will the indexes be composite headings, such as author/title? Or will the index contain just one type of field, i.e., author only, title only, or subject only? For example, NOTIS has composite headings or, I think, author, title, and date out on the side. Geac's headings simply display the heading itself--that's also true in the WLN system.

If you decide you need the composite index, what fields should be included and how much data should be used from each field?

If you have, say, an authority file that you want users to be able to scan, we're back to the old card catalog issue--is it going to be dictionary or divided? Penn State, I believe, has a dictionary index that's scannable. Most others--almost all other systems that I've seen--do not have that. You have to say you want the author portion, or the title portion, or the subject portion.

Comment: Anne Lipow tells a marvelous story about a user looking for Kropotkin. She came to the system and the system asked if it was an author, title, or subject search. Well, obviously, Kropotkin is an author. She put in "author" and got no books about Kropotkin. I find a lot of times it's difficult to explain to users that IBM is an author, when it seems to me, intuitively, it is a title.

Comment: Subject.

Comment: Well, almost anything you're looking for is the subject, isn't it? I mean, in a way. We use these words author, title, and subject, as though they have meaning outside of the tight little clique we're in. And

they have meaning, but they certainly don't have the meanings we assign to them.

Comment: We also use the term "series title" as though anybody could understand what that means.

Comment: The problem inherent in divided files is, how do you explain the division in such a way that all of your users understand the division? One technique that Columbia used in manual days, and other libraries have used, is a name file. I think they set up names as one file, and then everything else as the other--with titles and subjects in the other.

It also strikes me that titles are the subjects that authors assign, and "subjects" are the subjects that librarians assign. Somehow we feel that the ones we assign are more valid that the ones the authors of the works assign. Titles from an author's point of view are descriptions of what is inside, the author's attempt to describe in the way to sell it. When you get to situations where they are used that way, such as with dissertations--titles of dissertations have become very much less flowery since there have been Boolean systems to use them. There's a little chicken and the egg sort of thing here. I suspect that at the point that titles are used as major ways for the public to go in to things, authors are going to start putting a lot more precise titles on works, and not until then, whether ANSI says so or not. Legislating hasn't worked, but I think the market economy may well work.

Comment: I think we ought to have just two types of indexes, one called "it" and the other called "about." And if you want "it," you use the "it" index, and if you want "about" you use the "about" index.

Comment: I'd go further and say that the distinction between wanting "it" and wanting things about "it" is the distinction that people who do a lot of searching get. To a lot of users, there's not even that sense of distinction. Whether the thing is from--to use an example that gets really messy--a corporate approach, whether it's from IBM or about IBM, in some cases gets to be a kind of a fine distinction.

Comment: Not for IBM!

Comment: Part of the problem you face, however, if you're talking about a dictionary approach, is the fact that a person can be overwhelmed with information, more than what they really want. I don't know that people are that unfamiliar with author, title, subject, because they've been around the card catalog for one heck of a long time. They're also using them in various microform catalogs now, which are very often divided. In some cases, actually, physically different machines are used for author, title, subject. The big problem that I see in a dictionary catalog--and I would like to hear what--the experience of others is--is that you may just be overwhelmed with the thing telling you more than you really want to know.

Comment: It seems to me that the critical issue is to get the user into the scanning sequence on or close to where what they're looking for is. I'd be interested in hearing people's ideas and experiences as to the best way to do that.

Comment: That makes a great presumption--that is, that whoever it is who's searching the catalog does know in fact what he wants, and has some kind of description of it.

Comment: Even if he has some description of it, then there is the further assumption that that description is somehow right and that an accurate translation needs to be made, which may or may not be the case.

Someone was talking before about the translation of entry vocabulary into a vocabulary that the system proceeds with--which bothers me a great deal because now we're going to enter into a new set of vocabularies that we've got to translate. We've got document vocabulary, and user vocabulary, and card catalog vocabulary, and now we're going to have another new one. Too many translations.

One of the places, though, where scan search seems to me to be extremely interesting, and one that needs a lot of exploration, is in classification. We haven't talked too much about classification for a long time. We have the possibility of expressing relationships between things that exist in collections in a lot more powerful and interesting way than we have done up to now. We can assign more than call number to a book, for example, and express that in a classified kind of way--and I think that's going to take a good look at in the future. It provides a mechanism of transition from various vocabularies in ways that we haven't thought too much about for awhile.

Comment: I just wanted to follow this comment about classification, which makes me think of shelflist browsing. There are two possible assumptions. The searcher can be searching under a principle of substitutability versus non-substitutability. If they're looking for something, but cannot accept an alternate, then you're dealing with one kind of assumption. If they're looking for something, and they can accept an alternate, then you're dealing with another kind.

I personally don't know whether scan lists are good for exact match searching, whereas Boolean are better for "I'll take anything I can find in the area"--or some other functional distinction like that. Call number browsing is particularly powerful in the situation where the material that was first identified was in circulation, and a person wants to do exactly the same thing he or she always did on the shelf, which is to find a book next to it, which is substitutable.

I wonder if the differences between Boolean and scanning or set retrieval and scanning can be factored on this basis, the assumption of substitutability.

Comment: I'd hate to see the two different techniques that have been outlined so far linked in a singular way to two different kinds of searching. Even with the known items, the exact search for name, the user often doesn't know how to spell that name. Or with many authors with the same last name, they still may need the scan, even though they don't want a substitute. They only want that one work by that one author, only they don't remember the exact spelling of his name or his first name. It's still going to be useful for them to scan the main heading, especially with the name authority problem.

The one thing that hasn't been mentioned here--and I'll switch to a higher level on this issue--is the users have told us quite a bit through the surveys administered and analyzed ad infinitum over the past year, that we've heard so much about. They made it clear they want help in seeking related terms in the correct form; there's no doubt about that. They place a very high priority on that, whether they look at it as a problem or as suggested enhancements. I just wanted to throw that on the table--the survey data hasn't been mentioned yet this morning.

Comment: I want to return to the difference between how a user approaches a known-item search, or how catalogers do, and so on. As an ideal, if you want to pose a challenge, we would want to handle it similarly to how you would handle it in person. Somebody comes in, and you ask him, "What are you looking for?" They tell you, and you figure out, knowing what the system is all about, knowing what your techniques are all about, you focus on what is the likeliest source for this person.

This is where I would propose artificial intelligence techniques come in. After all, the types of searches that can be initiated in the online catalog are a very finite number. This is one area where very profitably in the near future one can reasonably expect that the system can figure out what type of search the user wishes to perform. The system probably cannot for a long time figure out topic or meaning in a subject search, that is all.

I propose that you give flexibility and allow somebody to type, "I want books on such and such a topic, published between this and that date, and so on." The system ought to be able to figure that out, much like some of the artificial intelligence applications commercially implemented with database management systems. The system can look in the index. If they find "New York," that occurs as a state entry and as a city entry. The system can engage the user in a dialogue using New York City and New York State by proposing this as an option. I think it's doable.

Comment: Knowledge-based systems.

Comment: Exactly.

Corey: One question that came in on scanning was, "Why do so few of the catalogs have the capability to display index terms?" If keyword terms are stored alphabetically, scanning of the keyword file may also be an aid to the user who wants to do a very fancy Boolean search but doesn't quite know what the keyword terminology is. Scanning has had a bad press. Everybody is ignoring it. So I just wanted to put it out on the table for a few minutes, and then we'll get to the real interesting topic of set retrieval.

Comment: One of the underlying assumptions that we've been making about scanning is that, in fact, you index in terms of something that's useful for the user. Suppose you, as a systems designer, decide to explore a technique for designing your indexes much like those of OCLC with their search keys. Scanning search keys from a user's perspective ain't too helpful.

Comment: We started out with a very abbreviated indexing style, and we found that the users simply didn't have enough information. They could not relate what they were looking for to the display. Consequently, I think, in most cases we doubled the amount of information we provided back in the screen.

Comment: I think it's basic to the whole thing that most of us are working with database management systems in which it is necessary to truncate, and in which it is necessary to have a key that is a fixed length. And I think it's a big problem for us in this particular area.

We get around it by the most brute-force method we can think of. We don't use any database management system that is provided by anybody else. We made our own and ours, in fact, handles full-length keys. Full-length keys, variable-length from one right up to whatever. In order to handle our scanning capability, we needed this capability. Nobody does it, nothing in IBM does it except ours. I don't know of anyplace else that does it.

Comment: Maybe just a quick comment. I think there are techniques in using keyword files that can effectively take any varying-length character string and reduce it down to an access point type of number into another file, giving you, in fact, screen compression and fixed length access. So I think there are ways that you can allow variable length search arguments, but from the database management point of view, keep your indexes very small and always index fixed length data.

Question: And you scan it?

Comment: Yes. Sure.

Corey: The last several questions have pointed out my initial remark that you can't talk about searching without having some assumptions about the file structure. Very hard to separate.

Comment: Scanning brings back, it seems to me, into the online environment a lot of precoordinated techniques that we used in batch environments. I think some of the indexes in DOBIS are KWICed, and certainly a relational database has its own kind of precoordination. I think when we get into scanning, we want to think not only of scanning from head of string but providing, in a sense, an exploded scanning mechanism.

Question: I'd be curious to know how many of the systems that have separate indexes for author, title, and subject have them because the librarians thought they were needed or because, from a system design aspect, that was thought to be easier or more efficient in terms of machine resources. There seems to be sort of an initial assumption that it was done because the librarians wanted a divided catalog. But is it really that, or is it also a design issue, as well?

Comment: In our system, we found that the librarians requested that author, title, and subject indexes be separate. Although, when we started to implement it, we found that the subject treatment and author treatment were so identical. In fact, the same file, the same codes are used in treating authors and subjects so you can get control of vocabulary for searches on author and subjects from one file.

Comment: I don't view it as a design problem in terms of efficiency. I think it was done because the librarians thought they wanted it, and I think it's our intention to change because I don't think they want it any more.

Comment: I think it's very necessary, though, to be able to make some distinctions, at least in real time, not necessarily predivide the headings. Simple, effective means of handling some of the problems on this point were mentioned earlier. Whether they're residing in there in dictionary form or not, it still may be useful for many searchers to be able to say what parts of the record--and we have been indoctrinated culturally to know that the broad distinctions, at least author, title, subject, and maybe a few others--are in there. That can be done in real time. A simple intermediate display screen that ILS uses, and keyword searching, in particular. Before they allow you to scan a display of headings, you use simple guides telling you--it's with postings--you found so many in one column, by the author.

Search for Einstein, and before it dumps you into Einstein, you would have this problem of divided indexes or some identification of whether it's by or about or from the Einstein Medical College, for example. It's a corporate thing, a conference on Einstein. All those are possible, and you will find they're in most large databases near Einstein, which is my favorite choice.

It gives you a guide screen, and you can choose among those. You can choose all of them. Again, it also forces narrowing at that point, if you only want to know about Einstein or if you want to look for a corporate author. So you tell them. You, in fact, divide it at that point interactively. And, of course, whether they understand the librarians' or the

catalogers' labels for that, we don't know yet. I don't think they do, but we'll find out through very simple empirical research. With a system where you can just change the labels to match those postings, and say, not "personal author" or "corporate author," but "by its author," or "about him." You know, it says your name appears in these places, and you segment at that point. It can go on from there, which is a simple way to solve the problem.

Question: How often is that distinction useful to users, and to what percentage of users? Two points here. It's not clear to me that, to a lot of users, the distinction is important. If I want to do a paper on Einstein, I want to know what he wrote and what we wrote about him. But on the other level, for how many people is there enough stuff out there--how many authors where the distinction becomes important? If you retrieve five items, the distinction is unimportant.

Comment: Not just authors but a lot of other words may appear in more than one place, and you can choose all of the above. I don't know, why don't we give them that guidance, instead of assuming they know everything, and dumping them into a divided index system?

Comment: Having looked up your first book on Einstein, you could probably go to the bibliography, and then your next question is--how many of these books do my libraries have, and what is their availability? And in that case, you're starting with the title. Yet the scholar's done the work, you've got it, so you don't want to wade through the rest of that list on the system. I think there's still a requirement to be able to type in "this is the title, I know it's the title, do you have this book, and can I get it?"

Comment: Yeah, but that's a known item and I agreed completely with you. What I'm saying is when someone is not doing the single item search, the known item search, but doing a subject search, which is where this distinction comes into play, how often and in what percentage of the real cases in real life is this distinction of any value? I sense that it's of value to a very, very small percentage of people, real cases in searching, both for reasons of set size and user need.

Commment: I can't agree with you. I think there may be a very few cases, but those are all very important cases, you know, like Mark Twain. Mark Twain, you can end up with 100 responses, sort of indiscriminate, and then you're faced with a tremendous burden to sort them out.

Comment: Which then suggests as a follow on, which somebody suggested, that we only use this technique in those cases.

Corey: Well, let's go on with set retrieval. Scanning was the easy one.

SEARCH OPTIONS FOR SET RETRIEVAL

The enumeration of search options for set retrieval, which are described below, begins with the simplest options and proceeds to the more complex. Set retrieval searches may be controlled vocabulary or keyword searches, or a combination of both using Boolean logic. The simplest type of set retrieval is a controlled vocabulary search without Boolean logic.

Controlled Vocabulary Searches Without Boolean Logic

In all controlled vocabulary set retrieval searches, except possibly number searches, right hand truncation should be automatic, and the system should retrieve all records that match the search data up to the length of the input. In general, if the search argument matches more than one controlled vocabulary heading, the system should return the set of all headings that match. If the search argument yields a unique heading, the set of all bibliographic records connected to the heading should be retrieved.

Controlled vocabulary search should be possible using a variety of access points:

1. ID Number Search. Number searches will usually retrieve bibliographic records. ID number access points include:
   - o Call Number
   - o Library of Congress Card Number (LCCN)
   - o International Standard Book Number (ISBN)
   - o International Standard Serial Number (ISSN)
   - o National Authority Record Number
   - o Documents number, as specified by the library (e.g., Superintendent of Documents Number, U.N. Documents Number, NTIS Number, corporate report number)
   - o Local system number, as specified by the library (e.g., accession number, circulation label number)
   - o Bibliographic utility record number as specified by the library (e.g., OCLC Number, RLIN Number)

2. Author Search. A search by author should retrieve all author names (personal, corporate, conference) from the authority file that match the search argument. If only one author name matches the search argument, the system should proceed to retrieve all bibliographic records associated with the author.

3. Title Search. The system should retrieve all titles that match the search argument. If only one title matches, the search should retrieve all bibliographic records that have the title as regular title, title added entry, or series title.

4. Series Search. A series search should retrieve all matching series entries. If only one series matches the search argument, the system should retrieve all bibliographic records belonging to the series.

5. Subject Heading Search. A subject heading search should retrieve all matching subject headings unless there is only one, in which case the system can proceed to retrieve the associated bibliographic records.

Should a user enter an author, series, or subject heading that is related to an authorized author, series, or subject heading by means of a SEE cross-reference, the system should probably display the search argument along with the authorized headings and give the user an option to continue. If the search argument matched only one heading and that heading had a single SEE reference to an authorized heading, the system could, as an alternative, retrieve the bibliographic records linked to the authorized heading in the authority control file without requiring the user to re-enter the correct heading, but it must inform the user about the action.

The online catalog should not require the user to formulate and enter OCLC-like derived search keys. The user should only enter a string of characters or group of words. The system may, at its option, translate the user input string of characters into a search key, for retrieval purposes. In such cases, the system must verify records retrieved before display and filter out those that do match the search value.

Keyword Search

Keyword, term, or uncontrolled vocabulary searching first of all should enable a user to search single terms against the author, title, series, or subject indexes to the database. This search retrieves every heading, which contains the search term entered by the user, from the index requested. Keyword searching based on a single term (and thus involving no Boolean operations) is limited in value but certainly not worthless. Hildreth has referred to this case as "manufacturing a high performance automobile without a steering mechanism."[6] Hildreth is correct in the sense that keyword searching with Boolean logic is far superior, but quite a lot can be accomplished, especially in small databases, with single term searches. The Geac system currently has single term keyword searches as the only type of keyword search. The keyword search on OCLC's LS/2000 currently limits the user to the entry of a single term. LS/2000 then presents the user with a list of the types of fields (author, title, subject, etc.) where the term is used and shows the postings for each type of field. If the user selects a field that has a low number of postings, the search continues. If the user requests a field with a number of postings that is above a defined threshold value, the system will prompt the user for a second term and will perform a Boolean AND on the two terms. Thus, while Boolean logic is available in LS/2000, many searches can be completed with a single term and no Boolean logic. Furthermore, the Markey study of subject searching in card catalogs[7] indicates that users frequently begin a subject search with a single term as

their entry point into the card catalog, and then scan through the cards using the other terms they have in mind to find relevant material. An online catalog system that can retrieve all subject headings having the user's "entry" term (whether the term is the first word or not) and then allow the user to scan the set of retrieved headings can be a useful search technique.

Notwithstanding the modest usefulness of single term searches, the online catalog should also allow the user to enter more than one keyword at a time. In large databases, a single term will often retrieve far more records than the user has the patience to scan through. The simplest form of a multiple term search is one in which all the words apply to the same field. For example, the keywords would all be title keywords, or they would all be subject keywords. When the user enters more than one term, each term should be linked by an implicit Boolean "AND" operator to narrow the search.

The system should allow the library to indicate which fields shall be subject to free text searching, although personal names will present problems unless special algorithms are used for them. It is desirable that the system allow the library to request keyword indexing of all MARC fields, including the notes fields.

Any online catalog with keyword search capability should maintain a stoplist of selected words. Any stopwords entered by a user for a free text search should be ignored by the computer system. The user should be informed that the stopword(s) used have been ignored. The list of stopwords should be customized for the database of each library. If a user enters only words found in the stoplist, the system should inform the user that this has occurred, and suggest alternative search strategies.

The system should also allow for truncated terms. Unlike controlled vocabulary truncation, however, keyword truncation should not be automatic. The system should require an explicit truncation symbol, such as the "#," for each term to be truncated.

In the special case where the user is doing a title search and knows the work has a one word title, the user should be able to supply a special character to indicate this situation, and the system should then not retrieve works with multi-word titles.

Boolean Searching Beyond Keywords

The online catalog should allow the user to conduct searches involving more than one index. For example, a user should be able to enter an author's name and a subject heading or, say, a title and a subject heading. The system should permit the user to furnish more than one search argument for the same index. The most common need for this feature occurs when the user is able to furnish two author names for a work. The system should also allow controlled vocabulary and keyword data in a single search. A common example is to give an author name and some title words for the desired work. Searches that

involve more than one index must retrieve bibliographic records, rather than stopping with the retrieval of the matching headings. Thus a search for, say, author "Williams" and title "Pediatric Nursing" should not return a list of authors named Williams and a list of titles beginning with Pediatric Nursing but should return the set of all bibliographic records satisfying both arguments.

Whether or not these searches using multiple indexes and multiple search values require the user to know and use explicit Boolean operators (AND, OR, NOT) will depend on how the user-friendly portion of the online catalog is designed. For example, Figure 3 shows a screen from the Biblio-Techniques BLIS system, which does not require the user to type any Boolean operators. Nonetheless, complex Boolean searching done by more experienced searchers will probably be easier through explicit use of the operators. Online catalogs should thus allow explicit Boolean searching for sophisticated users. For explicit Boolean searches, additional operators, such as those defining adjacency and context, are desirable.

FIGURE 3: Multiple index searching in the BLIS system
BLIS On-LINE CATALOG - Model 3

Catalog Searches
(fill in information)
Title Words:
Author:
Subject:
Subject:
Series:

Increasing the Number of Records Found in a Search

The online catalog should provide help to the user for increasing the number of records found in a set retrieval search. The need to suggest search adjustments is especially acute when the search has found no matches. Depending on the nature of the search that failed, the system could suggest that the user:

- Truncate, or truncate further.
- Use fewer key words that are connected by AND.
- Use more key words connected by OR.
- Change indexes.
- Change to broader terminology (for subject searches).

If the failed search is a controlled vocabulary author, series, or subject search, the system could automatically switch to browse and display that portion of the authority file that is closest to the users search argument. As an alternative, it is desirable, when an unsuccessful search occurs, that the system display the closest possible matches (fuzzy matching) and allow the user to make a choice on how to continue. The user should be told that a fuzzy match has been tried.

When a keyword search using terms connected by AND retrieves zero hits, one frequent cause of this failure is a misspelled word. The misspelled word has no postings, and thus the intersection of the misspelled word with the other words results in zero hits. It would be very helpful in the case of any key word search resulting in zero hits for the system to check the postings for each term and report terms that have no postings.

If the user is well into a search and a search refinement caused the retrieved set to become too small, the system should allow the user to return to at least the previous search result. As another search option, the system could have the capability to return to any one of several previous search results.

Decreasing the Number of Records Found in a Search

The online catalog should have several ways to help the user reduce the number of records found in a set retrieval search. The system could suggest that the user:

- o Give more of the heading (less truncation).
- o Add more key words connected by AND or NOT.
- o Add another index connected by AND.
- o Change indexes.
- o Change to narrower terminology (for subject searches).

The system could switch to browse for controlled vocabulary searches that produce too many hits.

In addition to the above techniques, the system should allow the user to reduce the number of records by limiting search results by:

- o Publication date.
- o Type of material.
- o Language.
- o Publisher.

In order to save a needless consumption of computing resources, any of these set reduction techniques could be suggested after a response from the computer indicating that a large number of records (more than say 75?) have been found and before the system has retrieved and displayed any of them.

One further possibility is to let the user see extracted samples from the large set. One type of sample is to select every nth record where n would vary with the number of records found and would give the user about 20 records to look at. For example, if a search found 3,000 records, the system could allow the user to see every 150th record, which would make 20 records available. The Northwestern NOTIS system has a more sophisticated extraction technique, which avoids displaying records that fall in the middle of a corporate name or in the middle of a subject heading with several different subdivisions. By eliminating headings with the greatest number of characters in common, NOTIS in effect backs up to the beginning of the corporate name or subject heading so the user knows what the beginning subdivisions are.

If the user is farther into a search and a search refinement (such as a Boolean OR) made the retrieved set too large, the user should be allowed to return to the previous set.

Search Scoping -- another kind of set retrieval option

In addition to the *ad hoc* interactive search adjustments that must be tried when a set retrieval search finds too many records, the system should also allow *a priori* criteria to be established, which will apply to all searches until changed and will have the effect of reducing the size of the retrieved set. One word sometimes used for *a priori* search criteria is "scoping." Presumably this word is used to indicate that the "scope" of the search is to be reduced according to the pre-established criteria.

One need for scoping pertains to subject headings. If an online system contains subject headings from several sources, e.g., LC, NLM, and Sears, it would be highly desirable under some circumstances to limit responses to one of the types of headings. Medical librarians, for example, may wish that subject searches done at medical library terminals ignore matches that are based on LC headings.

A second scoping criterion of high potential use is location. Many users might like the system to limit its responses to items available in one library or group of libraries. This would be especially useful for large library systems that have highly specialized or remote branches. Location scoping is almost imperative for online union catalogs that are also intended to serve as each library's individual catalog.

More Advanced Set Retrieval Options

In addition to all of the above set retrieval options, there are a few more options that are found very infrequently in online catalogs. At this stage of online catalog development, these features have to be regarded as the "more advanced" ones because they are more difficult to implement.

Related Record Searches

In the online catalog, records may be related in several ways:

- o Authority records can be related to one another by SEE and SEE ALSO references.
- o An authority record can be related to several bibliographic records by virtue of the fact that the bibliographic records "contain" the authority heading.
- o Bibliographic records can be related to other bibliographic records by virtue of the fact that they share an access point in common.
- o A series bibliographic record and an analytic bibliographic record are related in that they both describe, albeit in different ways, the same bibliographic volume.

Through record linking techniques, the online catalog should support searches for all of the above related record cases. The online catalog should display SEE references immediately and show a flag for headings that have SEE ALSO references, but the searcher should have to give a command to view the SEE ALSO references associated with a flagged heading. The online catalog should allow the user to request easily the bibliographic records that use (are linked to) any of the displayed SEE or SEE ALSO references. From a display of bibliographic records, the online catalog should also support the retrieval of related bibliographic records without requiring the user to rekey any headings found in the original bibliographic records. The Biblio-Techniques system has the HEADINGS command, which extracts all authority controlled headings from one or more bibliographic records and numbers them. The user can request other bibliographic records that have any of the headings simply by typing the appropriate line numbers.

The word "browse" is often used to describe a sequence of searches consisting of a mix of basic set retrieval searches, scan searches, and related record set retrievals. To Fox the essence of the browse search is the related record set retrieval,[8] and that is the meaning that has been reserved for the term "browse" in this paper. Using the meanings assigned in this paper, the WLN BROWSE command is not a browse but is a scan search, and the MELVYL BROWSE command is a keyword set retrieval search of all headings that contain the search word or words.

Set Manipulation

The user should be allowed to perform Boolean logic on saved sets or on a saved set and the active set (latest set retrieved). The Mankato State University PALS system and the Dartmouth online catalog allow users to combine sets by entering explicit Boolean commands. The Paper Chase system at Beth Israel Hospital in Boston disguises the Boolean operations by giving users an option at the bottom of some displays to "FIND PAPERS COMMON TO 2 OR MORE LISTS" and "INCLUDE (OR EXCLUDE) PAPERS FROM 2 OR MORE LISTS."[9] When a Boolean operation on sets produces an unwanted result, the user should be able

to issue a BACKUP command, which would return the user to the previous set. As an aid in helping the user decide on the previous set to select, it would be nice if the system were able to summarize the user's previous searches.

Proximity in Keyword Searches

Another online catalog option is to allow users to set explicitly a degree of proximity for words in a keyword search. The user should be able to vary the degree of proximity for any two words from none (the words must simply be in the record somewhere) to requiring that the words be directly adjacent. The user should also be able to state which word is to be first in any given word pair.

Series Searching

The system should allow users to identify a series and request a specific item in the series by volume number or year. The series could be found by either a controlled vocabulary or keyword search. The system should require the user to narrow the search to one series before allowing the user to request a specific volume. For series that are analyzed, the system should allow the user to search by either the series data or by access points from the analytic record. Either way, the system should respond with location information for the specific volume desired by the user.

There may be even further set retrieval options, but that's where I've run out. Let's throw the discussion wide open as to who is going to use these techniques when.

References

6. Charles R. Hildreth, "To Boolean or Not To Boolean?" Information Technology and Libraries 2 (September 1983): 235.

7. Markey, Process of Subject Searching. Examples provided throughout the report.

8. Fox, "Machine-Assisted Browsing," 79-80.

9. Gary L. Horowitz and Howard L. Bleigh, "Paper Chase: A Computer Program to Search the Medical Literature," in The New England Journal of Medicine 305 (October 15, 1981): 926.

Comment: You may have covered this implicitly, but if someone enters a subject heading phrase and the system says there's nothing, in some sense that is a very incomplete answer, particularly when the system could search an authority file and find out whether the heading exists. It's like asking a registration system who took Chemistry 113, and Chemistry 113 is not offered this spring; the system could easily tell the user that, but doesn't.

Comment: Some pretend to and don't do it very well. In the LCS system at Ohio State, the term you entered will appear as line 15 on the screen holding headings 11 to 20. Of the 30 headings it retrieves for you, you get 11 to 20 first. You get the second page, and then you go back to page one or on to page three. But the term you entered appears if, in fact, the case you mentioned is the case. It just leaves a blank out there in the postings column. We're doing a lot of controlled research now in protocol analysis, and it's confusing the hell out of us, because the system just leaves it blank. It doesn't say--even if they put a zero there, the users would like it better.

Comment: If zero really means there isn't anything in the library on what you want, then that's a very good result. The problem is when it means something different.

Comment: It usually means the latter, something different.

Comment: The worst case is getting some results that you can't figure out how you got.

Comment: I know there are things that you want to do to help the user, but I'm really worried about a system where I think that I'm doing one type of search, but the computer knows better, and it's going to help me, and it's going to do several other things for me. When you work with operating systems, there are some that are really nice because they do very simple things for you. It's the ones that can do 43 different things...I remember working in PL/1 and getting a 14-page description of how it would do arithmetic for me. It didn't think types matched. That can be a real problem.

Comment: That's been a concern of mine. The scenario we have been considering, for example, in our system, the subject, author, and corporate entries and so on are exact entry matches. That is, you can truncate them but it uses that string as a entry point into the index. If you happen to transpose a couple of words, or you don't get the right word to start with, then you're out of luck. You're not in the area of the index that would interest you.

If you do get no postings, it'd be nice to say, "We don't have any titles that start like that, or have this string of characters in this sequence that you entered. However, I'm now doing a Boolean search and combining all these keywords that you put in." Then return a set of weighted terms, if you will, back to the users, and present them with a list of titles that have some or all of those keywords in them.

Comment: What's wrong with doing the first step, and then asking the user, do you want to try it again, because you may have misspelled it, or do you want me to do this?

Comment: Yes, I think that's a good approach because one of my concerns in what I'm suggesting is doing these Boolean ANDs in the background, unbeknownst to the user. Are we suggesting that that's what we're doing? If that occurs very frequently, you're going to have a tremendous overhead on the system in superfluous searches, which can just eat you alive.

Corey: We might call this a class of secret searches. (laughter) Again, the one I know for sure is the one at Washington University Medical School. But it's a real tricky thing, to get into this.

Comment: That's PNAMBC searching.

Question: PNAMBC?

Comment: Yes, that's Pay No Attention to the Man Behind the Curtain. (laughter)

Comment: It's interesting; we were talking last night about the operating systems that have to face some of these problems. I know in the DEC 10, it had a very sophisticated file structure program. But one difficulty it has is that if you mistype a file name, so that you're searching for a file that doesn't exist, that is the single highest overhead that it does. Because it now has to search its entire tree. And it does it all for you. That's a problem, because 20 or 30% of entries represent somebody leaving out a letter. We don't want our default searches or erroneous searches to be the single most expensive thing we can do in the system.

Comment: I've followed up on several thousand user searches, and looked at exactly what they put into the log, and used my expert ability as a card cataloger to see if I could outperform them. In a surprising number of cases, I can.

An example of the searches that I have proposed is that when somebody wants a subject, instead of asking, "Do you want a subject in authority control" or "Do you want a keyword," or whatever, the computer would just start in and look for an exact match on the subject authority file. If it finds it, show it. If it doesn't find it, then do a keyword search on the subject authority file, and show him. If it doesn't find that, then do a

keyword search on the title. All of it transparent to the user, and coming up with some kind of final subject list.

I've looked at thousands of these searches, and over and over again after you get into them and work on them a couple of weeks, you see the user doing something stupid. You say, you know, you're on the wrong track. You know this guy is doing the wrong thing, and he's going to come up with nothing in the end. Wouldn't it be nice to have the computer interrupt him at that point and say that you're on the wrong track, try a different approach?

Comment: There was a study that Stanford did similar to that. A certain number of reference searches were done using the card catalog and then using an automated system.

Comment: Yes, that was done in our reference department. Essentially that was to prove that using RLIN we made more productive use of our reference librarians than handling things in a card catalog.

Question: Do you remember the outcome as far as the hits were concerned?

Comment: It was basically a timing study, as I recall. We were not comparing the accuracy of results. Basically for single items. We weren't comparing--we were just comparing time, and in all cases you could find the item in the card catalog, and you could find the item in RLIN.

Question: Somebody puts in a title word search, or whatever, and somebody is looking for a corpus of material on a given topic, my assumption has always been that it's up to the system to present everything--not just something useful, but everything that exists. Is that a sort of universal assumption or not?

Comment: You can ask the user: do you want this approach or not? In other words, the machine can ask if that is one of the assumptions. I tend to agree with it more than I don't. But the dialogue can handle that creatively, and have both options at that point. Should I operate under this assumption or that? Let the user tell you. You've got some of that in sight. There are hardware and software constraints, but I'd like to forget those for a while, because 10 years will be here pretty soon.

Comment: The other thing, though, is as you get into related terms, as the profile linkages and these relationships become broader and broader, and they have more of a structure in place, I don't think you can say I want to show everything, because you get in to a...In fact, even going into a subject-oriented index, if you happen to hit on something that happens to have a bunch of cross-references, you can all of a sudden find yourself with a list that is much larger than you need, because you really did know what you were going after. So having a system automatically say: yes, I want everything related...

Comment: Oh, I didn't mean all at once.

Comment: I think this issue is tied to levels of interface. If you only have one level of interface, you have to take a posture of one kind or another and do it one way. Presumably do it the best way you can to help the average user.

If you have an interface that recognizes that there are different types of users that approach the online catalog, full spectrum, the first-time newcomer to the experienced searcher, then you can indeed, at each level, carefully consider what your philosophy should be and what you do. Ultimately, at the fullest level of the interface, you can give full control to the user. At the beginning level, you probably want to be selective.

Corey: Yes, that's the big question that comes in the end after we enumerate all these options. The question is which ones come into play when. I don't mean when, you know, the third step of the search. I mean when depending on the skills of the user, if that can be detected.

Comment: It would seem pretty clear that each of these techniques has its uses, but no one of them is sufficient.

Corey: Right. That's again why I said I didn't have the answers for you because this is a complex one.

Let's go on to the next problem of set retrieval. That's the other one we're all familiar with: decreasing the size of the set when too many things have been found. There are some parallels here to the first problem of finding nothing.

The first option again, it seems to me, is somehow to inform the user in such a way that they would try to search again, perhaps providing more information. If they just put in author Jones, maybe they know the author's first name so they could supply more of the search argument. If they are using a keyword search, they could supply another term that might tend to reduce it.

Another option that we're seeing in some systems if the set is very large is providing a summary of the set. I don't know if there is consistent terminology for that. I think in the Northwestern system you call it the guide card technique, where you do a search and you get 1,000 records back. It's not quite like providing every nth record; the algorithm is a little different from that.

Comment: Well, it was basically predicated on the idea of guide cards in the card catalog. It's automatically determined by what you retrieve in terms of where it gets truncated.

Corey: You condense it down so it only takes one screen. Then the user can kind of narrow the subset out that they're interested in--the part of the set. Does LC do that, do you know?

Comment: They do in the MUMS system, yes.

Corey: Okay. Is anybody else doing that? That's the only two I know of.

Comment: The guide may be based on some other value such as type of field.

Comment: What the CITE system does, for example, is if you went in with a name and the system finds that the number of records is under the given threshold that it can handle, it automatically tells you how many items and you can browse the records if you want to. If the threshold is exceeded, then the system automatically goes into a neighbor display mode and you choose the item or items that you want in the alphabetical display. So it's the internal threshold that guides you. Because that's, in fact, what a trained searcher would do. If they go in and say author Smith, the system would blow up and say sorry, too many or whatever. Then they would use a neighbor command to look at neighboring terms in the alphabetic index.

Comment: In our SPIRES implementation we can display every nth record. So, if you get 5,000 records, you can display every tenth one if you want, every fifth one or whatever.

Question: Does that get used? How many times do people say: well now, take me through these 500 records one at a time. Does anybody have an idea of the amount of access, the amount of activity that kind of a command gets?

Comment: I'd just guess it's infrequent.

Comment: I think it's the screen that comes up that leads them that way. Then I think you're getting somewhere. But just that command to do it, I imagine most people would find that--well, I found our experience was quite --one of those nice little whistles but nobody ever blows it.

Comment: We have some data to answer that question, but I can't remember it. (laughter) It's there in our analysis of the transaction logs from the University of Syracuse system. The situation is roughly the same. You're not prompting the scan. Generally, that's all you can do--scan short records. We also know that it's not infrequent for the user to get 150 or more. This way you just ask for five at a time, I believe it is. We have some very good data on exactly what they use.

Comment: They run like every fifth? Every nth record? They pick a kind of random sampling through the list?

Comment: Yes, they kind of do that in a way.

Comment: We originally tried the every nth record approach and it's got a lot of problems. It can turn out to be very misleading to the users for reasons that would take a little time to explain.

Comment: There's a strategy that I use when I take a topic, and a lot of other people also use it. If I'm trying to find the importance of material in a given area, for example, if I'm looking for something, if I'm doing a paper on the Great Depression, I don't know what the subject headings are that relate to that. If I can find just one bibliographic record and get the tracings, if I can find something by Galbraith, you know, and look through those, get the tracings, and then reissue the search. Systems are starting to do that sort of semiautomatically.

Comment: Just two things about looking at every nth item. It does make sense, I think, under certain circumstances. In the CITE system, the users are told that the system automatically displays closest match items. It's not an exact Boolean-type search. So, in that situation, the person knows that 500 records were retrieved, but the closest matching items come on top. As you go down the line, periodically the system asks: "Do you want to continue with this kind of stuff?" In that kind of a situation, we recognize that the option of giving the choice to look at every tenth or fiftieth item may be valued so you have an idea of how far down you might still seek with things. So, in the successor system we are working on, which is not yet operational, CITEHILL, that is a choice. But it is because we know that the user knows that it's a diminishing return type of display of records.

As far as user-friendly interfaces are concerned, I think the best guide is to take a look at the native capabilities of the system, what you really can do on a given system. Then take a look at how experienced searchers would handle a particular problem like searching for author names with truncation. From that type of behavior and techniques, build into your user-friendly interface the type of choices or type of approach where the system automatically emulates, if you will, much of what the competent trained searcher would do, given that your system can do certain things.

Comment: I invented a terminal some people might be interested in. It had a little lever on it. You push it all the way forward and it would skip every--well, you know, like 20 percent of the total file on every display. You back up a little bit, about every 500 entries. Then if you'd back up a little bit more, it would display every one. Then, if you push down it would go back down through the alphabet. I did invent it for this purpose. I didn't think people ought to have to type when they use the card catalog. But it's cheap and easy to do. The only problem with it is you have to design your whole file structure and everything to take advantage of that.

Comment: I would also put a gas pedal. (laughter)

Question: I know that the AND is of supreme importance among the Boolean operators. Relative to that importance, how important is the OR and the NOT, and what priority should be given to those? What priority are the users giving to them in their actual searches, for example?

Comment: Again, there's a lot of data on that. OCLC has some and Mike Cooper has just reported on a study of a huge mass of searches in the last (November 1983) issue of JASIS--a study of a huge mass of searches on ELHILL by trained searchers. We have a lot of data where the author is ORed in the Minnesota system. To summarize it, the OR option is chosen fairly infreqently. I again don't have the exact percentages, printings, or population per system, but that's the base line. OR is not used frequently. It's used very infrequently by users of online catalogs where it's available and they can choose that option, explicitly choose it. I have a copy of Mike Cooper's article where it gives some hard data on that huge sample of searching ELHILL on how frequently they use the OR right down to whether it's one, two, how many times. There's one number--in fact, in Mike Cooper's transactional analysis of ELHILL, which relates to this but a little more generally. It's like 48% of the searches in this ELHILL sample were single word searches. That just blew me away. Forty-eight percent of searches in this huge ELHILL sample, trained searchers and everything, were single word searches.

Question: What about the NOT?

Comment: I don't recall. I think AND is the biggie and OR is the--we've got to get sufficient data for those.

Comment: The problem with OR is that users don't understand that when they say they want the primary results from the election for New Hampshire AND Vermont, they really mean OR. Somehow or the other we have to teach them to use OR, I think.

Comment: I think that's an important point. About five percent of our Boolean searches are ORs or NOTs, but my guess would be that that would go up considerably if people understood; but even the librarians don't understand.

Comment: You raised an issue early on and we haven't followed it through. That was the one on fuzzy retrieval, fuzzy sets. Implicit in what 90% of the people around the table have been saying is the idea of precise sets. It seems to me that fuzzy set retrieval is appropriate for users who have fuzzy requests, and the subject request is likely to be a fuzzy request--as opposed to this edition of Shakespeare's that. A fuzzy set response, I guess, is like: listen to what I mean, not to what I say. When the user is going against the system, it's not clear to me that the precise response is the best response. And I think that the findings of the studies CLR has sponsored over the last couple of years suggest that the majority of people

are looking--their intention is subject. A significant portion is actual subject searches, and some portion of what we would term known item searches turn out to be the Galbraith entry into the subject search. It may suggest that we've been barking up the wrong tree, and that Salton was right, and Hillman was right, and Tom Doszkocs is riding up the right street. That for the majority of uses of our catalog, fuzzy responses are better than precise responses.

Question: Has anybody had any experience with using phonetic type of access for that fuzzy type of search? Do we have any kind of experience with that?

Question: Phonetic?

Comment: Phonetic, yes. Running through a phonetic type algorithm.

Comment: You mean SOUNDEX or something like that? It certainly exacerbates the problem that we talked about earlier.

Comment: A McAllister search.

Comment: Spelling approximation probably sounds better than it really is. Something less than 2% of all searches that would have failed are made successful because of spelling correction.

Comment: If it is true that most of the searches are subject searches--which now is supported by data--if it is true that even in the known item searches quite often users have imprecise or incomplete information, then I think it's very useful to consider alternatives like the best match or closest match type search. That implies that if there is an item that is an exact match, it will automatically come on top. So, you don't miss out on anything, but you also provide the user other things to look at, which often are the things they actually turn out to want.

Comment: This relates to that example. I don't have the faintest idea what the subject heading is for the Great Depression. It might be the Depression of 1929, and I put in the Great Depression, and that produces absolutely nothing. But there is a subject heading there that has Depression in it, and I'd like to know what it is.

EPILOGUE

Figures 3 and 4 summarize search options for scan and set retrieval searching, respectively. Both tables show the types of search data a user may enter and the types of records that may be retrieved. In several systems, the first response is not a list of authority or bibliographic records but a list of index records. In some cases these indexes are composites of several fields that are truncated to fit into an 80-character CRT screen line. For example, the index might contain 30 characters of author followed by 46 characters of title and 4 characters of imprint date. In other systems, the index records contain only one field, such as a title field, and are truncated at 80 characters. In Figures 3 and 4 the term "index" refers to online catalogs that first respond with index records.

Most types of searches in most online catalogs have only one target, e.g., the bibliographic file or the index to the bibliographic file. The library user does not have a choice of target records. There are a few exceptions. Some systems, on a few searches, will allow the user to decide the target file. For instance, on a set retrieval type of search by author name, the Biblio-Techniques system will let the user choose between the authority and bibliographic fields. That is, the user can retrieve the set of all authors whose names are "Weber, D" (or fuller) or retrieve the set of all books that have "Weber, D" (or fuller) as an author. Since the latter option allows users to bypass the authority file and thus possibly miss helpful cross-references, it is doubtful whether this type of option is beneficial to the library public. It is probably useful for library staff who have a correct name and want to eliminate steps when searching for bibliographic records. At any rate, in most searches in most systems the target records retrieved for a given type of search are determined by the system designers and not by the user at the time of the search. The apparent choice shown for records retrieved in Figures 3 and 4 is usually not a choice for the searcher.

Some of the sub-options listed in Figures 3 and 4 are options that the user can control for each search request and some are options that are decided by the catalog designers. For example, the decision to display the count of bibliographic records associated with each authority record is made by the catalog designers, but the decision to scan backwards is made by the user (assuming the designers have put this feature in the online catalog). Of the options that can be selected by the user, some are default options and will occur unless the user suppresses them. Others require extra data entry on the part of the user. These sub-options, when combined with the variety of search arguments and the choice of performing scan or set retrieval searches, present the online catalog user with a large array of search techniques. It is an array that is in fact so large that the challenge becomes one of presenting a simplified subset to the user who neither needs nor has the patience to learn about the total search capabilities of an online catalog system.

FIGURE 3: Scan Search Summary

Search Arguments -- Records Retrieved	Sub-options
Author -- authority or index.	Implicit truncation.
Title -- bibliographic or index.	Forwards and backwards.
Series -- authority or index.	SEE refs displayed.
Subject -- authority or index.	SEE ALSOs indicated.
Call number -- holdings and bibliographic.	Multiple subject authorities (LC, NLM, etc.).
Keywords -- keywords or phrases.	Display count of bib records.
	Multiple shelflists (call number scan).
	Choice of sort order (bibliographic scan).

FIGURE 4: Set Retrieval Summary

Search Arguments -- Records Retrieved	Sub-options
1. ID number.	
LCCN -- bibliographic.	Implicit truncation.
ISBN, ISSN -- bibliographic.	
Gov't document number -- bibliographic.	
Call number -- holdings and bibliographic.	
RLIN -- OCLC number or bibliographic.	
Authority number -- authority.	
Piece number -- holdings and bibliographic.	
2. Controlled vocabulary.	
Author -- authority, bibliographic or index.	Implicit truncation.
Title -- bibliographic or index.	Explicit Boolean.
Series -- authority, bibliographic or index.	Limit results by date, etc.
Subject -- authority, bibliographic or index.	Automatic SEE reference follow-up.
3. Keyword.	
Author -- authority or bibliographic.	Explicit truncation.
Title -- bibliographic.	Implicit Boolean AND.
Series -- authority or bibliographic.	Explicit Boolean AND, OR, NOT.
Subject -- authority or bibliographic.	Limit results by date, etc.
Notes -- bibliographic.	Completeness flag.
	Word proximity.
4. Combinations of the above -- bibliographic.	Explicit Boolean.
5. All of the above.	Summary extraction for large sets.
	Backup to previous set.
	Backup several sets.
	Boolean on saved sets.
	Related record searches.
	Revise search by returning to search argument.
	Scoping.

Selecting the best kinds of search options for online catalogs is an issue that is far from settled. Not only is the question of which options to have available in the system unsettled, but the corollary issue of which options should be presented at various search stages to novice searchers and searchers of medium skill is even further from a clear solution. For this Baltimore meeting, CLR asked participants to submit questions in advance for each program topic. Many of the participants' questions on search retrieval options did not get discussed at the meeting because there was not enough time to cover them. Nevertheless, the questions are important and need to be addressed. The written questions submitted by participants on the topic of search retrieval options but not discussed at the meeting are the following:

Questions About Scanning Options

1. What is the nature of the records that will be scanned?

   Specifically:

   a. Will the records contain full headings without regard to length or will the records be truncated headings of, say, a length that fits conveniently on one line of a standard 80-character CRT?

   b. Will the records be a composite index of headings such as author/title or an index that contains just one type of data, i.e., author only, title only, subject only?

   c. If indexes are composites, what fields are included and what amount of data should be used from each field?

2. Will the author, title, and subject records be separate (as in a divided catalog) or merged (as in a dictionary catalog)?

3. What sort of "filing rules" will be used to order the records? Will filing rules for authors be different than rules for subjects?

4. How should SEE ALSO cross-references be made known to the user who is scanning?

5. If the user's search argument does not exactly match an index (as will frequently happen), where in the index should the response begin?

6. Should systems offer a fast forward and a fast back or is some other method preferable when the user finds he is not alphabetically close to the desired heading?

7. What attempts should be made to "normalize" the indexes beyond converting lowercase letters to capitals? Specifically:

a. punctuation.

b. diacritics, apostrophe, hyphen, special characters.

c. double dashes in subject headings.

8. If a system has a keyword index, should users be allowed to scan it?

9. Why do so few online catalogs have a capability to display index terms that are alphabetically adjacent?

10. When a user has selected a heading and has requested the associated bibliographic records, in what order should the records be returned to the terminal? In other words, will it be possible to scan bibliographic records in some intelligible way? If the search was conducted on a series, would the order of retrieved bibliographic records be by volume number instead of main entry? Should the order of records retrieved from a subject search be different still?

11. If headings were rotated so all words occurred in the first position, would that affect the answers to question 10?

12. To what extent is scanning of controlled vocabulary indexes, by itself, adequate for online catalogs? Is scanning adequate as the way to begin all searches, with reliance on Boolean logic and set retrieval only for searches that do not succeed using a scan display? Are scan searches more successful for some types of searches (personal author?, title?) than for others (subject?, government name?, conferences?)?

Questions About Set Retrieval Options

1. Granted that a broad search that produces a large number of matches is both expensive of computer resources and of little value to the user, how big does the retrieved set have to be before the system:

   a. completely refuses to do (or continue) the search and instead suggests alternative search techniques?

   b. warns the user, but still gives the user the option to request the search to continue?

2. In what sort order should the retrieved set be displayed?

   a. if the records are authority records, they should probably be in alphabetical order, but where do subdivisions of subjects and corporate bodies file? E.g., U.S. Lawn Tennis Association.

b. if the records are bibliographic records what should the sort order be, and does the type of search make a difference? That is, should the order of bibliographic records retrieved by a

   i. subject search be by publication data?
   ii. author search be by main entry?
   iii. series search be by volume number?

c. if the records are holdings records, should the sort order be by location, copy number, piece number?

3. Related to questions 1 and 2 immediately above, how many records can the system sort and still give good response time? Or for large sets, should the system present bibliographic records unsorted? Or better yet, is there a way to retrieve bibliographic records in sort order?

4. Should the system permit left truncation or "don't care" characters in the middle of words?

5. Should normalization of indexes used for set retrieval be any different from normalization rules for scan indexes? What normalization rules should apply to keywords?

6. What algorithms should be used for fuzzy matching? Phonetic indexes? Is fuzzy matching worth the cost?

7. How important to library patrons is access by such keys as LC card number, ISBN/ISSN, OCLC/RLIN system number, manufacturer's number?

Questions About Keywords and Boolean Logic

1. "Boolean Logic" is a term that is understood by a comparatively small percentage of the people who will use online catalogs. How extensively, therefore, should explicitly defined Boolean searching be incorporated into the design strategy of a public access online catalog? Should implicit Boolean searching be incorporated into the design strategy, when one considers that Boolean searching is generally more expensive in computer resource usage as compared to single term/phrase searching?

2. It is generally considered desirable to have at least some type of Boolean search capability in an online catalog. Should there be several types of Boolean searches available to the user? Can Boolean searching be made less complex for the casual user?

3. After much debate, title keyword searching has come to be generally regarded as a necessary online catalog search facility option. Keyword seaching of subject headings, however, is still under active discussion, particularly as regards systems that support linked authority files. The

fact that arguments against keyword searching of subject headings frequently invoke the "could never do it in the card catalog" reasoning suggests that keyword searching of subject headings may soon be taken for granted too. What are the pros and cons of keyword searching of subject headings, particularly in light of the emergence of systems that support linked authority files? Are designer choices going to be forced anyway by what the public seems to regard to be the easiest way to conduct subject searching?

4. Does the presence of keyword search and Boolean capabilities reduce the need for authority control?

5. If keyword searches go direct to the bibliographic file, there is the chance that the search terms will belong to an unauthorized heading that has a SEE reference to another heading, but the user will never know. Should keyword searches for authors, series, and subjects always be directed to the authority file first?

6. How will increasing public computer literacy engendered by PCs in schools and homes change user demands on the systems, specifically for sophisticated Boolean capabilities?

7. Given large databases (1,000,000 records), how do we make Boolean searches fast yet effective?

8. How can users of an online catalog be trained to perform efficient Boolean searches?

9. Does the need of library patrons for a Boolean search capability justify the cost?

10. How much does the end-user actually utilize Boolean OR or NOT operators when conducting his own searches?

11. What techniques are available to make keyword searches of personal names work efficiently? Is it worth the bother?

12. What degree of proximity should be required of words in keyword searches? Must all the terms be in the same field? Same type of field? For example, should a title keyword search be limited to the 245 field or the 245 field plus other title fields? If author names are keyworded, is it essential that the system retrieve only records in which the supplied name words are in the same field? How important is it to avoid the case where a book has joint authors of, say, John Smith and Mary Brown and the system retrieves the book in response to a keyword search on Mary Smith? The same question applies to subject retrieval. If a user does a subject keyword search on, for example, "quality management" and a book has the subjects, "Quality of work life--United States" and "Employees' representation in management--United States," should the system avoid retrieving

the book? On the other hand, if subject keywords must be in the same subject heading, what relevant records will be missed?

Questions About Limiting The Retrieved Set

1. To what extent, beyond the usual date, form, and language limits, should direct access be provided by fixed field data? (I've been asked to provide limits by government publication, conference proceeding, and festschrift fields.)

Questions About Search Scoping

1. Location scoping would obviously be applied to the retrieval of holdings records, i.e., holdings would only be shown for the locations requested. But should location scoping apply to bibliographic records so that titles in the database would not be retrieved if they were not located at the specified location? Should location scoping be applied to the authority file?

2. How should scanning work if location scoping is in effect? Should the system not tell the user about authors, series, or subjects if these headings are not linked to items at the specified location?

Questions About Finding Related Records

1. Please provide a general discussion of the search facility options including the requirements to take the results of a keyword search (rank the frequency of assigned subject headings) and automatically link to the one or two most frequently occurring subject headings (similar to the Dialog EXPAND command).

2. Should the online catalog have the same syndetic structure as the card catalog with which the user is presumably familiar?

General Questions

1. Which special methods (e.g., truncation, keyword, adjacency, Boolean) should be implicitly invoked or introduced with special online guidance to different classes of users?

2. Why are there so few systems that display term postings for index terms? Is it computing resources, development, or imagination that's needed?

3. Is there any evidence whether keyword or subfield searching is preferable for subject searches?

4. Even experienced, trained professional searchers make relatively little use of sophisticated searching capabilities--why does everyone want them for public use?

5. What search facility options are essential? Which are frills? Are they different depending on the users and/or database?

6. To what extent can/should users be expected to understand or learn the implications of various searching options, and to what extent should they be expected to effectively control the choice of options?

7. Should an online information access system offer more than one type of search facility (such as Boolean, keyword, etc.)? If so, what types would be most beneficial, when should they be used, and when shouldn't they? How can the system itself detect when the user is using the wrong type of facility, or is that possible?

Although the large number of topics and the tight schedule at the CLR-sponsored meeting in Baltimore did not permit systems designers the time to discuss the relative importance of many search options, it is unlikely that additional time would have brought agreement or answers on the best role of many of the possible search options. As several participants said many times at the meeting, these questions really have to be answered by the users of online catalogs. The questions cannot be answered by systems designers, nor can they be answered by library staff, not even by reference librarians who work most closely with library users.

Empirical data are needed to provide system designers with guidance for search options. One source of data is the transaction logs that are recorded by most online catalog systems. Studies of transaction logs are just beginning.[10] More log studies are definitely needed. Another source of information will be obtrusive studies of online catalog users similar to the earlier obtrusive studies of card catalog users. The recent studies sponsored by the Council on Library Resources[11-18] are just a beginning. The results of those studies uncovered very general kinds of problems of searching online catalogs, such as the fact that users had difficulty knowing what to do when too much or too little was retrieved in response to a search.[19] The studies did not find out which techniques used by which systems worked best to help the user continue the search to a successful conclusion.

Several times at the meeting the issue of standards arose, and each time the participants concluded that it was still too early to set standards. The group readily agreed that the lack of hard data about the efficacy of online catalog features made it impractical to set standards at this time. In the meantime, systems designers have no choice but to rely on their own intuitions, the ideas of the library staff, and suggestions from library users willing to offer them. Every online catalog has, in its own unique form of implementation, answered the above questions concerning search retrieval options and many other questions of a similar nature. The answers are quite

different among the systems. Until the needed empirical data are available, online catalogs will undoubtedly continue to exhibit a variety of options for search retrieval.

References

10. John E. Tolle, Current Utilization of Online Catalogs: Transaction Log Analysis, Volume 1, Final report to the Council on Library Resources. Report No. OCLC/OPR/RR-83/2 (Dublin, Ohio: OCLC Office of Research, 1983).

11. Neal K. Kaske and Nancy P. Sanders, A Comprehensive Study of Online Public Access Catalogs: An Overview and Application of Findings Volume 3, Final report to the Council on Library Resources. Report No. OCLC/OPR/RR-83/4 (Dublin, Ohio: OCLC Office of Research, 1983).

12. Rosemary Anderson, Victoria A. Reich, Pamela Roper Wagner, and Robert Zich, Library of Congress Online Public Catalog Users Survey, A report to the Council on Library Resources (Washington, D.C.: Library of Congress, 1982).

13. Karen Markey, Online Catalog Use: Results of Surveys and Focus Group Interviews in Several Libraries Volume 2, Final report to the Council on Library Resources. Report No. OCLC/OPR/RR-83/3 (Dublin, Ohio: OCLC Office of Research, 1983).

14. Public Online Catalogs and Research Libraries, Final report to the Council on Library Resources (Stanford, Calif.: The Research Libraries Group, 1982).

15. J. Matthews and Associates, A Study of Six Online Public Access Catalogs: A Review of Findings, Final report to the Council on Library Resources (Grass Valley, Calif.: J. Matthews and Associates, 1982).

16. Users Look at Online Catalogs: Results of a National Survey of Users and Non-Users of Online Public Access Catalogs, Final report to the Council on Library Resources (Berkeley: University of California, Division of Library Automation and Library Research and Analysis Group, 1982).

17. Ray R. Larson, Users Look at Online Catalogs: Part 2. Interacting With Online Catalogs Final Report to the Council on Library Resources (Berkeley: University of California, Division of Library Automation and Library Studies and Research Division, 1983).

18. University of California Users Look at MELVYL: Results of a Survey of Users of the University of California Prototype Online Catalog, Final Report to the Council on Library Resources (Berkeley: University of California, Division of Library Automation and Library Studies and Research Division, 1983).

19. Davis B. McCarn, comp. and ed., Online Catalogs: Requirements, Characteristics and Costs, Report of a conference sponsored by the Council on Library Resources, December 14-16, 1982 (Washington. D.C.: Council on Library Resources, 1983), 19.

IV. USER FEEDBACK IN THE DESIGN PROCESS

Charles Hildreth

There are two possible interpretations of the topic of user feedback in the online information retrieval context. The first interpretation involves monitoring the use of online catalogs, then analyzing the captured data offline and reporting the results of systematic analyses to systems designers. The second interpretation involves monitoring user activity on an online catalog, and giving real-time feedback to the user during the actual search session to assist in the formulation of the query or interpretation of search results.

The first approach has been the focus of the recent CLR-sponsored online catalog studies, most especially what has come to be known as "transaction log analyses." This approach has proven very useful in describing what large populations of users actually do when searching online catalogs. The second approach assumes a fair understanding of online catalog use, and so builds on the results of the analyses undertaken in the first approach. Without denying the very great value of the second approach, this paper will focus on the first, and especially transaction log analysis.

Implied in our interpretation of user feedback in the design process is the ability of the system to incorporate the results of the analyses. If a system cannot adapt, there is little real value to feedback. This point is brought home in a portion of a recent article by Ray Larson and Vicki Graham of the MELVYL project at the University of California.

> In order for an information system, such as an online catalog, to provide effective service to its intended users, two things are required: first, the system must be flexible enough to change over time in accordance with the needs of its users, and second, some way must be found to determine those needs, providing feedback to the system design and development process.
>
> In designing public access online catalogs, this process of feedback and refinement is even more critical than in other types of information systems, both because of the relative "youth" of online catalogs (with all of the unknowns that implies) and the wide range in experience of potential users.[1]

Similar observations were expressed by David Penniman and Wayne Dominick in their paper, "Monitoring and Evaluation of On-Line Information System Usage."

> If information systems are to be truly responsive to users' needs, change itself must be considered in the process of systems design. Systems design can no longer be viewed as a one-time effort resulting in a static design that is unchanging for the operational life of the system. Rather, it must be predicated on the principle that change of some sort is inevitable and, in fact, desirable.[2]

Perspectives of the System Development Process

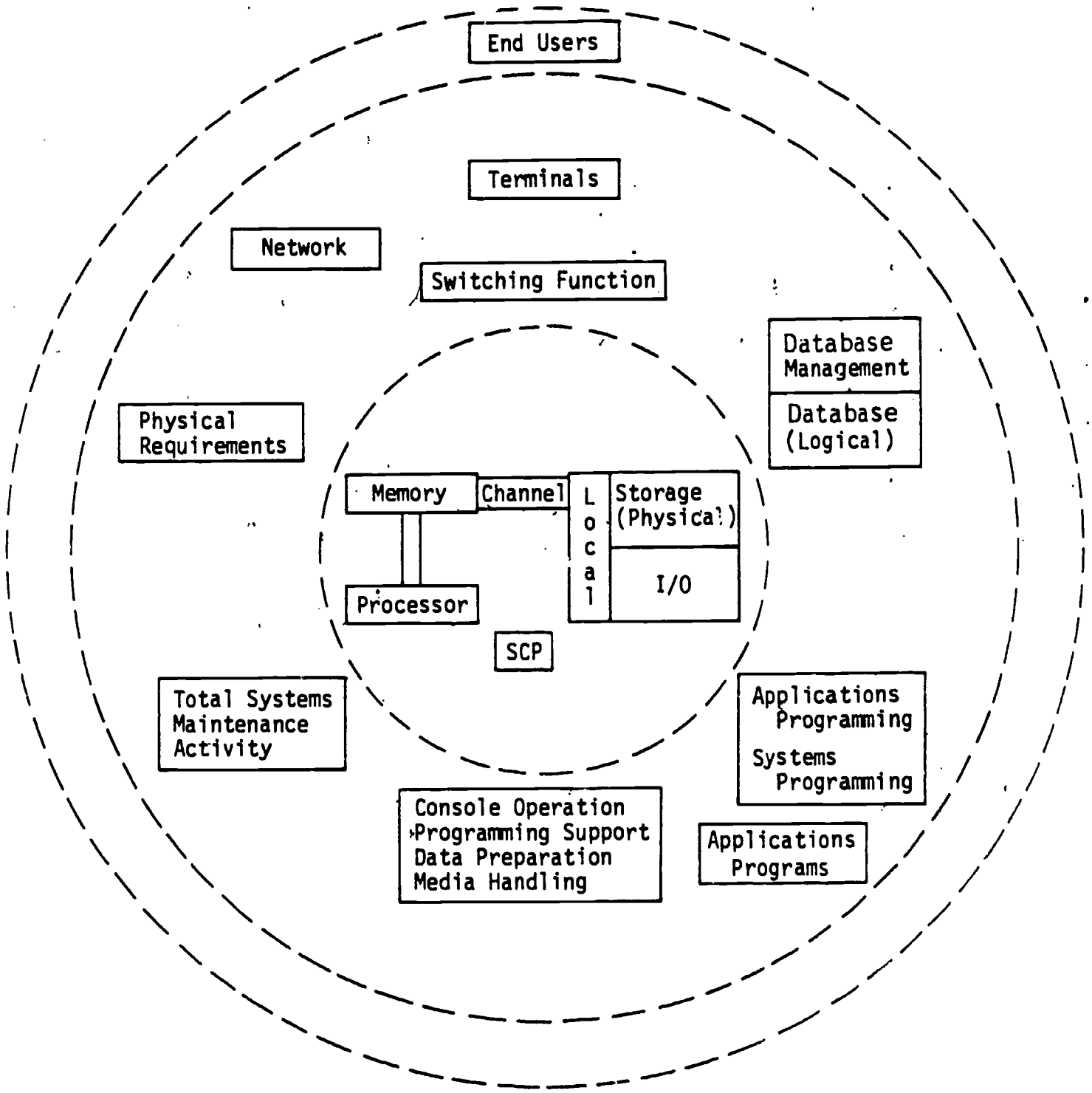
A few years ago, one of our research assistants on an online catalog project, a graduate student at OSU, came up with an interesting way of looking at the evolution of online systems. He said that we had moved from user-indifferent systems to user-hostile systems, and then to user-accommodating systems. While I suggested to him that there were also systems that could be described as user-friendly, his observation rings true to a very large extent.

The conventional computer systems design process has been, indeed, very user-indifferent. Figure 1 taken from Milton Marcus' 1977 paper, "Concerning the Nature of Information Processing End Use," describes the conventional view of the systems development process. This approach prevailed until about 1980. At the core is the central hardware and basic systems software. Around it is an array of peripheral devices, software tools, and applications modules. In Marcus' own words, the end-user, "the outermost band is as yet undefined and is included thus far only in certain unique situations."[3]

The conventional perspective is that we create the system and deliver it to the end-user, who is then the recipient of its valuable, marvelous features. It is simply the old sense of delivering the product after having developed and produced it. I think we have seen a shift though in recent years.

Beginning in the late seventies and gaining acceptance in the early eighties, the user came to be considered one of the critical, integral components of the system, not just at the periphery, the recipient of the system but not a part of the system, as Marcus depicted. The user is considered a critical component of the system and not merely the recipient, if for no other reason than that the successful working of the whole system depends upon how well the user manipulates the system. Thus, there is a new interest in how the user behaves when using the system, and some interest in shaping that behavior from the point of view of what is best for the system.

FIGURE 1: System Components and the End User: The Conventional Perspective (MARCUS, 1977)



Recently, some writers have even talked about using behaviorist techniques, for example, operant conditioning, to mold the behavior of the user so that the system will be used more efficiently and effectively. There has clearly been a shift from the conventional perspective depicted by Marcus, that is, from a system-oriented design to a user-oriented design. Today's online systems are commonly lauded as "user friendly" or "user cordial."

While observing the variety of attributes such products incorporate, I have been grappling with defining the "user-oriented" system design process. Figure 2, taken from Online Public Access Catalogs: The User Interface,[4] represents another model of the design process. The traditional areas that Marcus identifies, the hardware and software, remain important. The user is much more in evidence in this model, however. The focus of the diagram, as shown by the shading, is on the dialogue between the system and the user. The centrality of design emphasis has shifted from hardware/software to the interaction between the user and the system. Conventional systems that the user is familiar with, the traditional catalog and cataloging records, must also be considered in the design of interactive systems.

Another view of the major considerations in designing online retrieval systems is provided by Figure 3 from Pauline Atherton Cochrane.[5] This graphic outlines the major variables that system designers must take into account when designing user-oriented systems.

Changes in the Design Process

We have a fuller perspective on matters now, and I think our perspectives are definitely more user-oriented. We have a richer notion of what all of the important ingredients are, but we still do not know how to mix them appropriately to produce well-designed systems oriented to various classes of users. We are now aware and sympathetic to this rich, mysterious mix of elements and components required in the design of online systems. But has there been a change in the system design process that incorporates these richer and newer perspectives? As recently as 1980, Penniman and Dominick thought perhaps not. They could not see in their review of the literature at that point that we really had created a different design process that would do justice to newer, richer perspectives on the whole problem area and the entire environment of the system in the real world.

> At the same time that information systems were going through such radical change, their design process remained relatively stable. A conscientious systems designer first gathered data from existing or potential users, designed the system, tested a prototype, implemented the full-scale system and then waited for user reaction that might

FIGURE 2: The User Interface in Online Information Retrieval
(Hildreth, 1982)

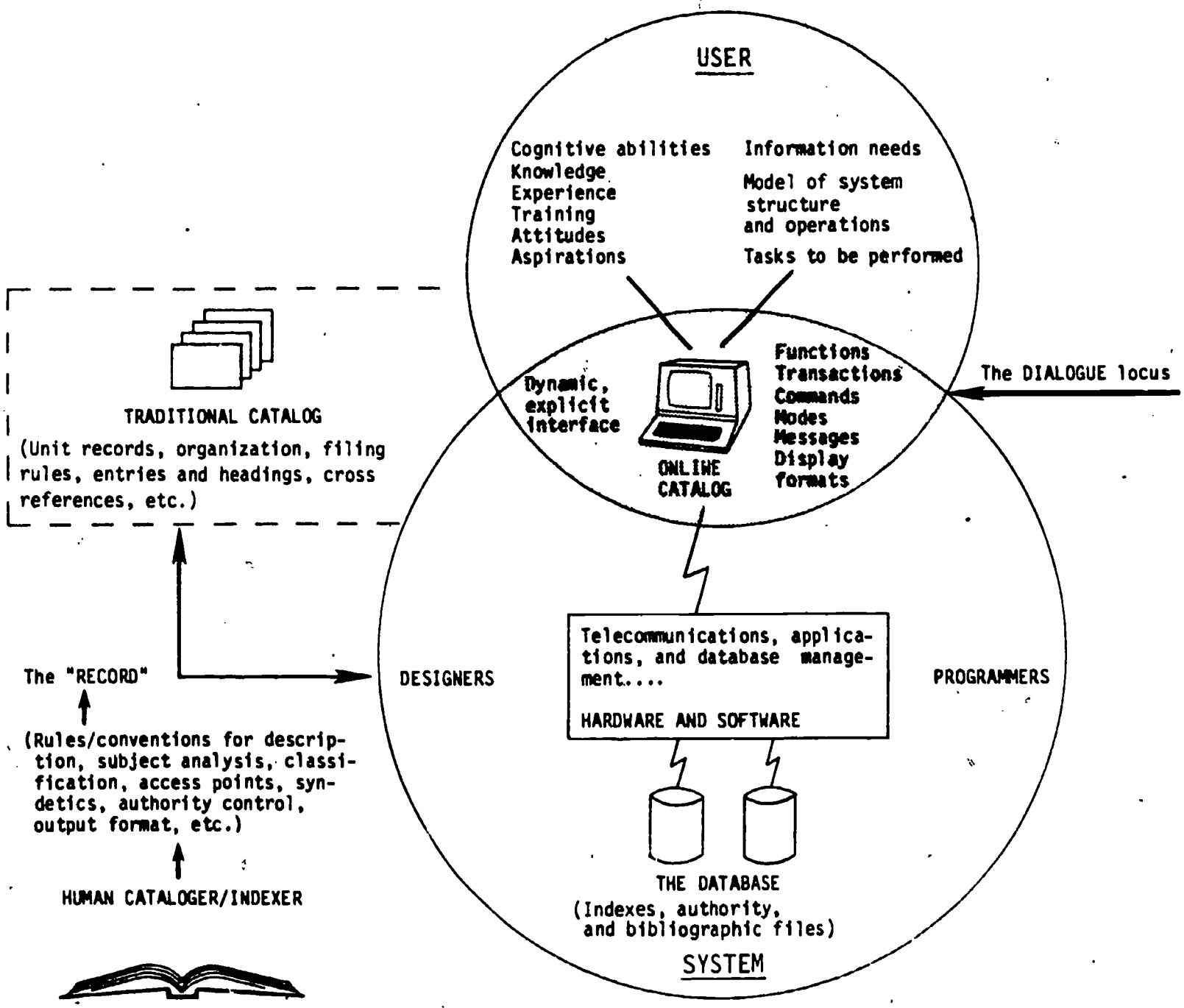
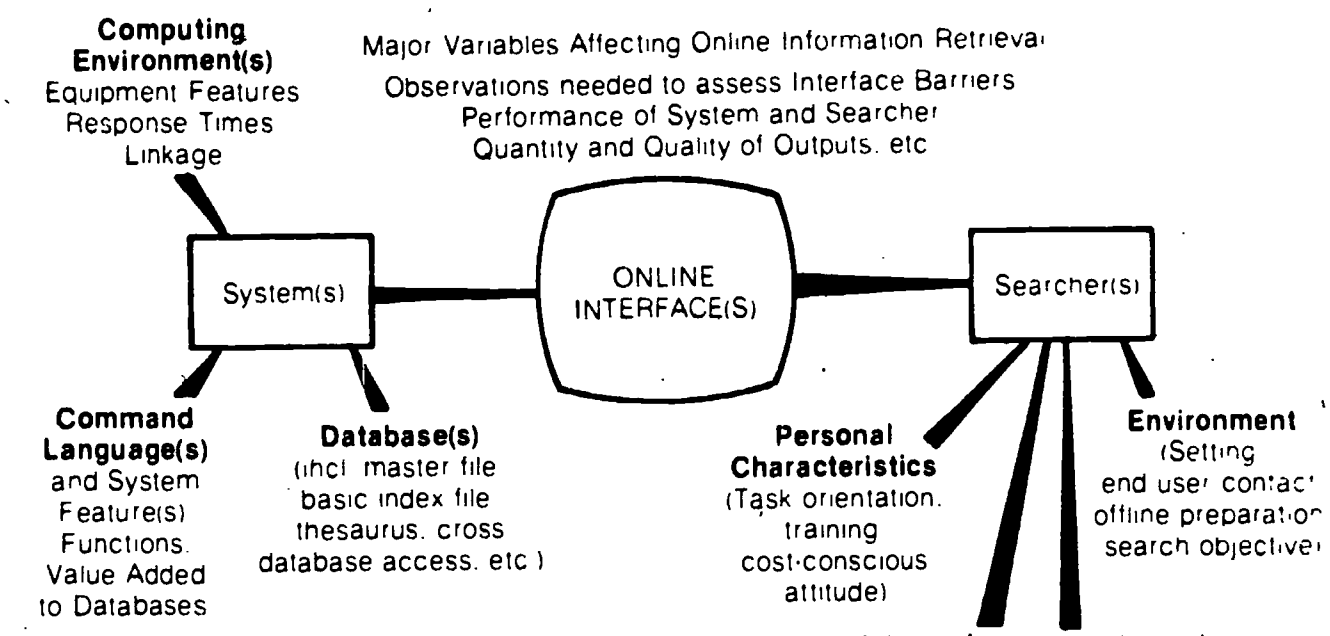


FIGURE 3: Major Variables Affecting Online Information Retrieval (COCHRANE, 1981)



> lead to system modifications (provided the design was able to accommodate change). While lip-service always was given to "user-oriented systems," the user to whom the system was oriented existed primarily in the designer's mind and tended to be more systems-oriented than the actual user group.[6]

Figures 4-8 illustrate what I see as an evolutionary change in the system design process over the last few years. Through the late seventies, the design process remained essentially as Penniman and Dominick described it, as shown in Figure 4. The design process consisted of system designers merely designing products to be developed and tested in a one-way, linear manner. The design process ended when the system was manufactured and delivered to the users.

Somewhere along the way, we introduced users, as a design problem, into our view of the process, as shown in Figure 5. You have still got the retrieval system out there, produced by the designers. Of course, users are using the system now, which was not necessarily taken into account before. Users are actually using it, interacting with it. I have added one other thing here, the little box down at the bottom, user requirements. At some point, it clearly came into play. Designers did not systematically interview users, but they started thinking about users.

Figure 6 indicates an awareness of various classes of users of the system. User trainers, such as the library reference staff, are coming into the picture now as system evaluators. These users interface directly with the end-users and the system itself, as well as the system designers. The design process is getting noticeably richer.

Figure 7 represents a very recent evolutionary stage. In addition to using the system interactively, users are being monitored and surveyed (online in some cases) to provide valuable feedback directly to system designers, within a very short time span.

The final graphic in the series, Figure 8, represents our current stage of design understanding. You will note that intuition has been replaced by the data fed back to system designers from various classes of users, user trainers, and library planners. This does not mean that intuition is no longer valued highly, but it is much less important in the design process. We now have a lot of user feedback mechanisms, multiple channels of communication.

We have acquired an entirely new understanding of what we mean by "system." The system is not simply the end product, what is delivered to the user; rather, the system is all of the components and feedback channels depicted in Figure 8. This goes back, of course, to general systems theory, the systems theory of communication. The system includes all of these factors

FIGURE 4

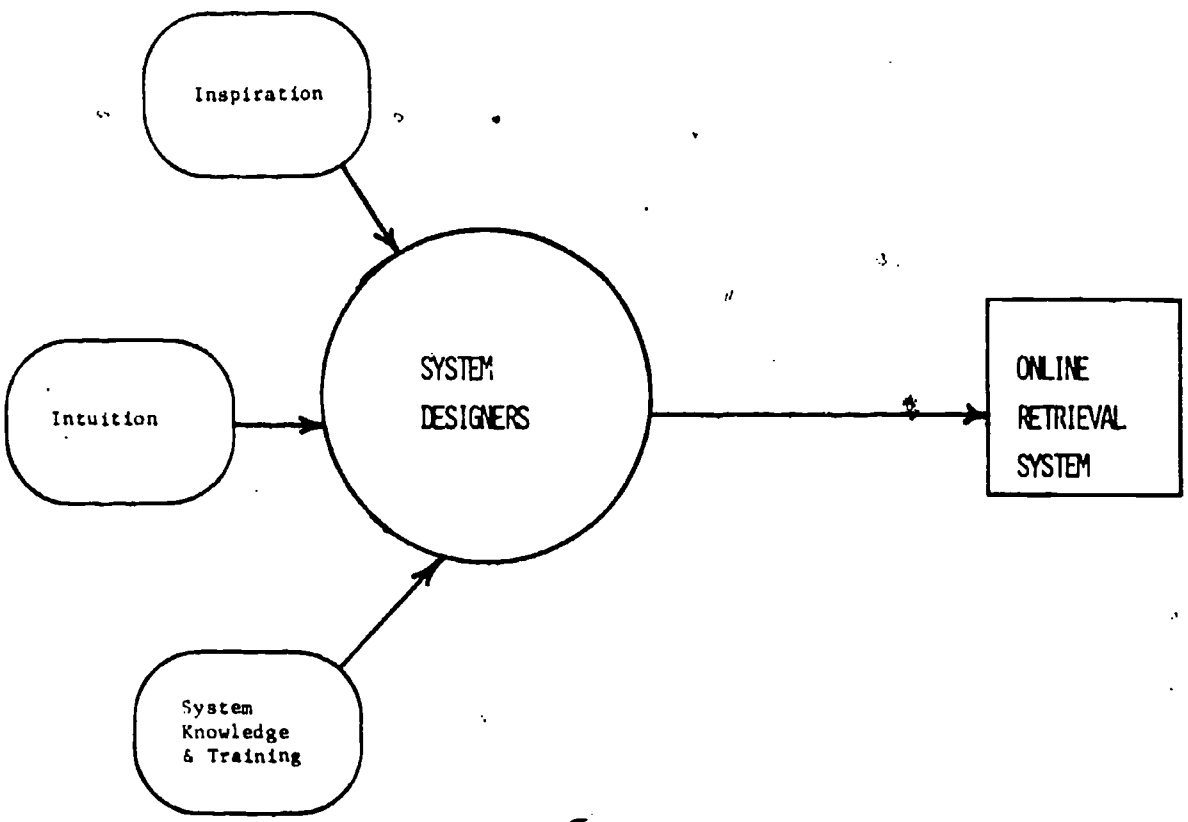


FIGURE 5

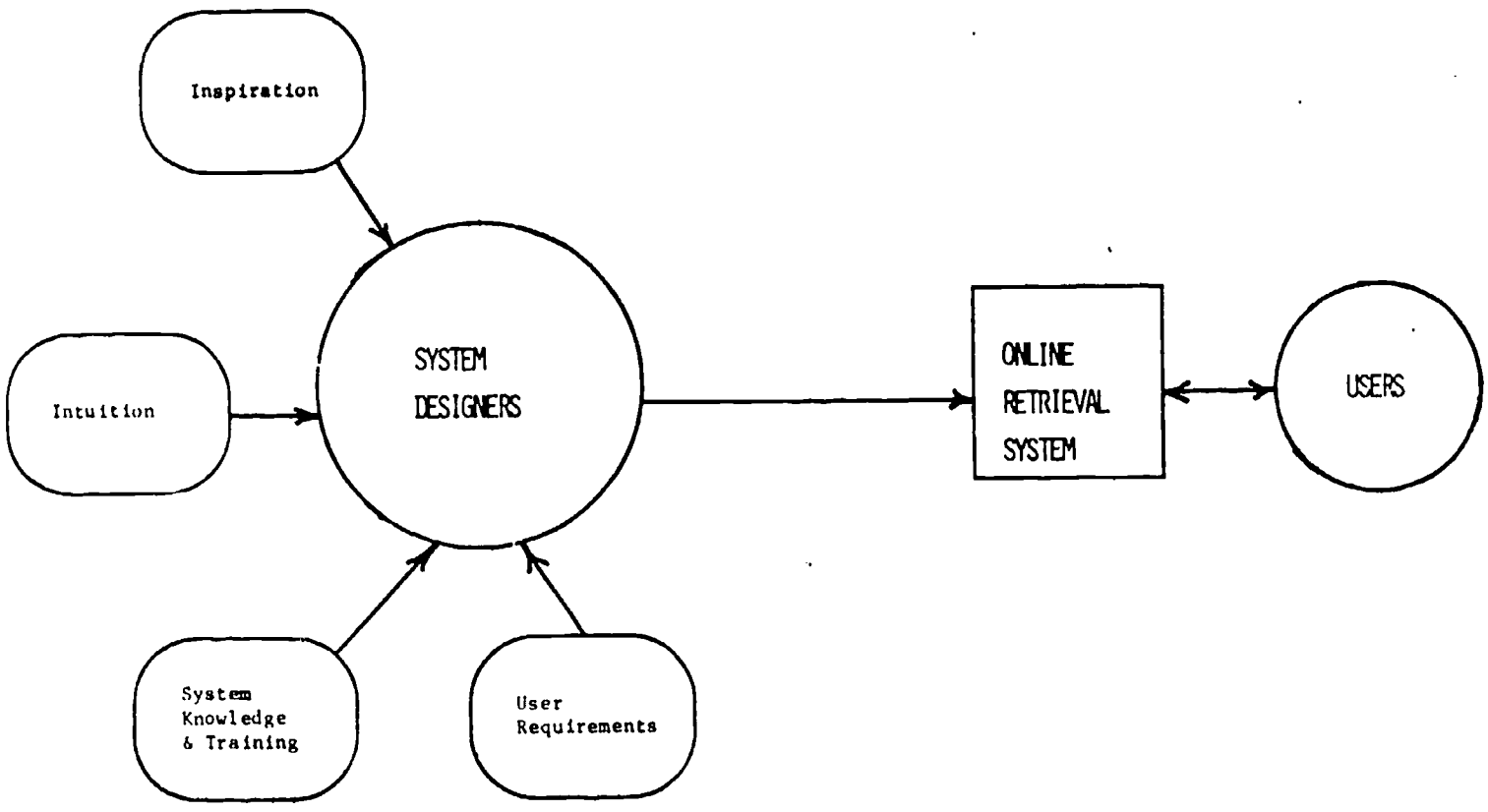


FIGURE 6

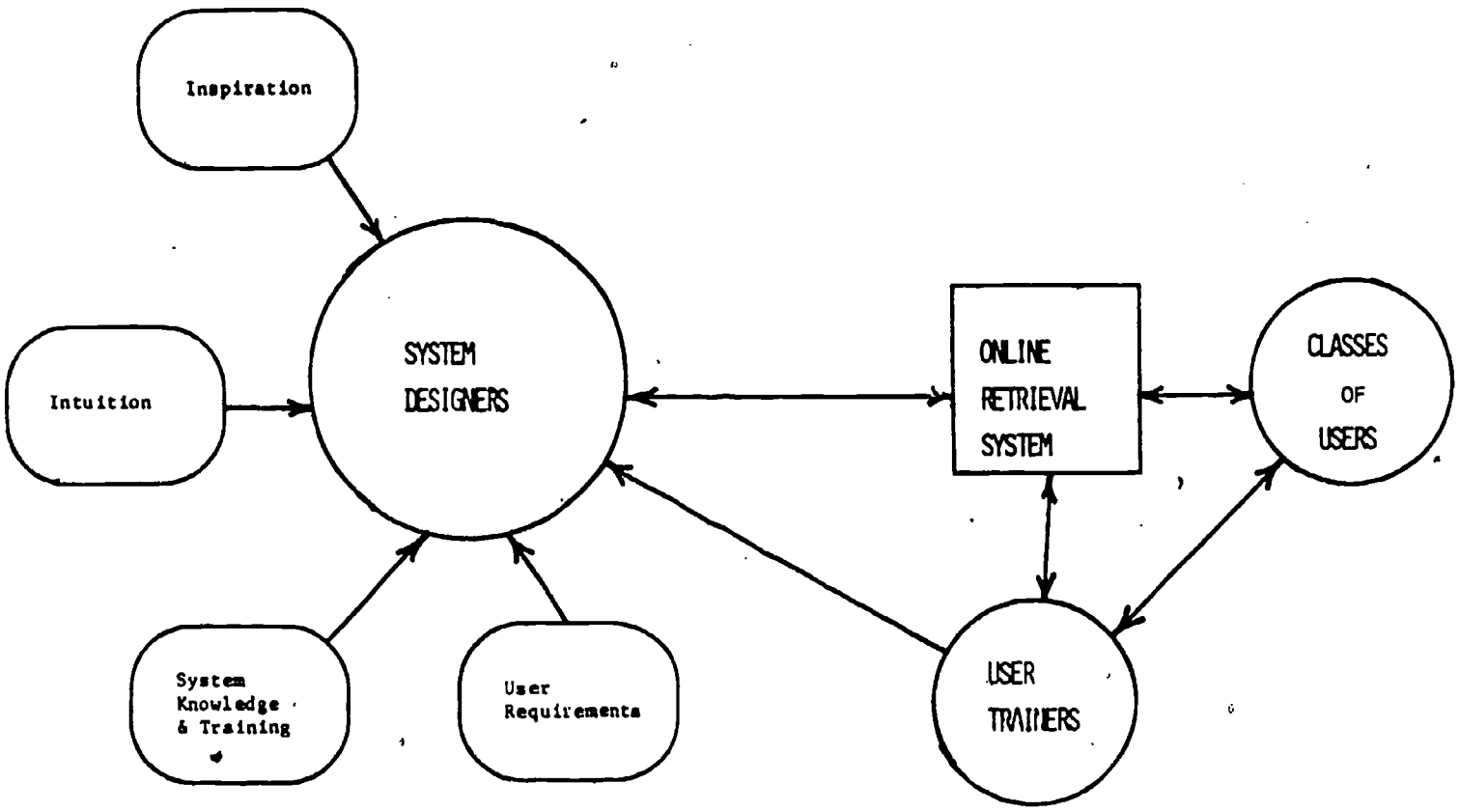


FIGURE 7

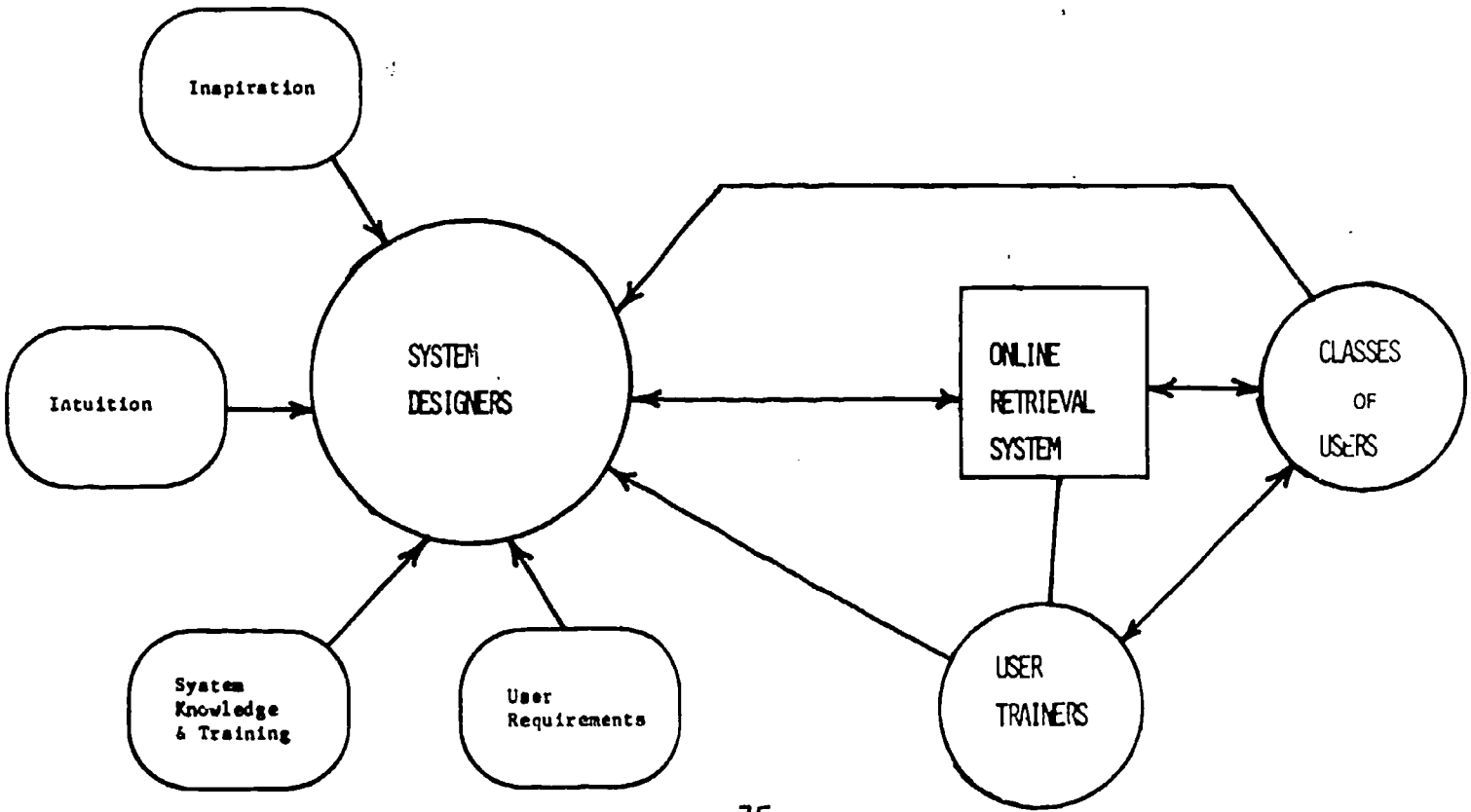
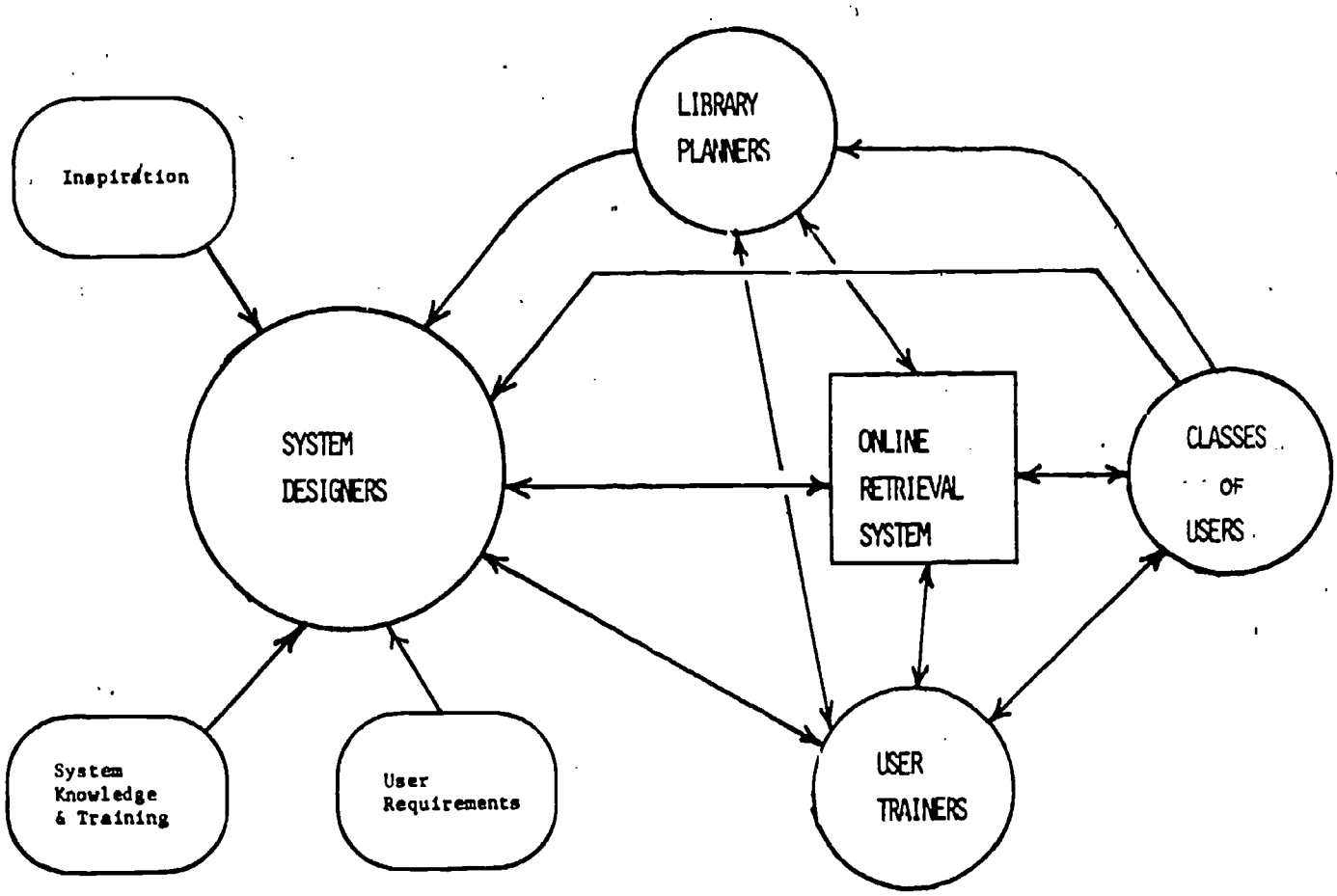


FIGURE 8



in the design process. And it is process. The design process is part of the system. The system is not the end result of the process, but the entire design/development process itself, including the product. the users, and the feedback mechanisms.

The system as design process includes the feedback mechanisms. Just as the end-user came into importance in the last two or three years, not just as an after-the-fact design consideration, now we have to add one more stage. The feedback mechanisms incorporated in the process are just as important, and should be dealt with just as systematically as the end-user's concerns, needs, and requirements. We have seen a shift from the role of the user in the design process (as recipient or problem variable) to the role of user feedback mechanisms, i.e., user behavior/attitude monitoring and evaluation procedures.

How do we keep the system, in this larger sense, user-oriented? We must be careful that we do not allow feedback mechanisms and monitoring evaluations to become mere ivory tower exercises. Feedback mechanisms must be viewed as a part of the system, that is, part of the total design process. Carefully designed and administered user feedback mechanisms provide the systematic basis for a user-oriented system product.

This new, expanded, holistic perspective of the system (design process + system + user + feedback + evaluation) suggests multiple performance objectives to be defined in pursuit of improving the performance of the information system. Improving the performance of the user. Improving the performance of the designers. Improving the performance of the trainers. Improving the performance of the feedback mechanisms themselves.

Ben Shneiderman, of the Computer Science Department of the University of Maryland, one of the better-known writers on human factors in computer systems, has put together a useful set of steps for the design process, or what he calls human-engineered, user-oriented interactive systems (Figure 9).[7] Some of the design process steps are well known. He lists eight of them. The first step is collecting information. Steps two through six are what we usually think of as the design process: performance specifications, technical specifications, prototyping, software development, choosing hardware. The last two steps of the eight are worth emphasizing. Nurturing the user community is number seven. That is just good customer relations, client relations. He lists some things under that, including an online gripe or suggestion box in addition to things like telephone consultants, hotlines, and newsletters, users' groups, and all that sort of thing under number seven. It is a major step, but not the final one, in the design process.

The final step is preparing an evolutionary plan. For example, design for easy refinement or repair. Improve error handling. Measure user performance regularly. Carry out experiments. Sample feedback from users through questionnaires and interviews. In other words, systematic research is recommended as part and parcel of the good design process, not something we may do if we have time.

The Role of Research

We often confuse the ways of thinking about research and its value in our practical day-to-day environment. There are the traditional research approaches with which we are all familiar (see Figure 10). Sometimes we confuse them when we are thinking about what research we should be undertaking, and its role in the design and development processes. These approaches include investigative, experimental, evaluative, and predictive research. These are just useful distinctions, not always that separated in actual research activities.

FIGURE 9

Design Process for Interactive Systems
(Shneiderman, 1980)

1. Collect information.

   - organize design team
   - obtain management participation
   - submit written questionnaires to users at all levels
   - conduct live interviews where possible
   - read practical and academic literature
   - speak with users and designers of similar systems
   - estimate costs and cost/benefit
   - prepare schedule with observable milestones

2. Design semantic structures.

   - define goals and establish a hierarchy of requirements
   - consider task flow sequence alternatives
   - organize operations into transaction units
   - create application domain data structures
   - develop application domain operators
   - specify privacy, security, and integrity constraints
   - obtain agreement on semantic design

3. Design syntactic structures.

   - compare alternative display formats
   - create syntax for operators
   - prepare system response formats
   - develop error diagnostics
   - specify response time requirements
   - plan user aids and help facilities
   - evaluate design specifications and revise where necessary
   - carry out paper and pencil experimental test

4. Specify physical devices.

   - choose hard or soft copy device
   - specify keyboard layout
   - select audio, graphics, or peripheral devices
   - establish requirements for communications lines
   - consider work environment (noise, lighting, etc.)

FIGURE 9 (cont'd)

5. Develop software.

   - produce top-down modular design
   - consider modifiability, generality and portability
   - emphasize reliability and maintainability
   - provide extensive system documentation
   - conduct thorough test

6. Devise implementation plan.

   - assure user involvement at every stage
   - write and field test training manuals
   - implement a training subsystem or simulator
   - provide adequate training and consultation
   - apply spiral/layered/phased approach to implementation
   - aim to please the users

7. Nurture the user community.

   - provide on-site telephone or consultants
   - offer online consultant
   - develop online "gripe" command or suggestion box
   - make user news available online
   - publish newsletter for users
   - organize user group meetings for discussion
   - respond to user suggestions for improvements

8. Prepare evolutionary plan.

   - design for easy refinement or repair
   - measure user performance regularly
   - improve error handling
   - carry out experiments
   - sample feedback from users by questionnaires and interviews

FIGURE 10

RESEARCH APPROACHES

INVESTIGATIVE	(EXPLORATORY: KEY VARIABLE IDENTIFICATION, HYPOTHESIS GENERATION, ETC.)
EXPERIMENTAL	(CONTROLLED, HYPOTHESIS VALIDATION)
EVALUATIVE	(DIAGNOSTICS, PERFORMANCE TESTING, ETC.)
PREDICTIVE	(SIMULATION, MODELING, HEURISTICS, ETC.)

We undertake investigative or exploratory research to identify variables that affect user attitudes/perceptions of the system, their behavior, their actual use of it. I am not sure we have done all that we need to do in that area yet, that is, hypothesis generation, or, more simply, generating questions that need further research and study.

Then, of course, there is experimental and evaluative research. System designers have been pretty strong evaluators, especially if you are talking about system performance and the use of system monitoring devices and feedback mechanisms. They have been less systematic and perhaps less dedicated to the design of monitoring and evaluation techniques for evaluating user peformance or, beyond this, the integrated, combined user-system performance. Last, there is the predictive approach, which has been directed to retrieval system design only very recently.

These research approaches can be useful, mixed and matched appropriately, depending on what we want to know, what the questions are, and what our resources are, of course. But, to get a little bit more concrete, we need to know the objectives of the research. Of course, there are different study aims (see Figure 11). You might want to discover a variety of things, confirm a hunch, evaluate a new feature. Certain techniques or study methodologies are more appropriate for one of these aims than another (see Figure 12). You have to be very careful to decide what your aims and objectives are before you jump into any particular kind of research using a specific methodology, whether online monitoring of search activity, transaction analysis, just eyeballing the user's protocol (regenerated on a hard-copy print of a user's online session), interviewing users applying any of a number of interview techniques, or administering questionnaires and surveys. We should keep in mind that there is a variety of study methodologies that may be used to systematize, improve, or refine those user feedback mechanisms depicted earlier. In a recent article, Cochrane and Markey review "several methods to study library patrons' reactions to OPACs (online catalogs), user experiences and behavior, system features, patron use of and reactions to system features, and system performance."[8] The authors warn that not all research questions will be answered by a single methodology. They then go on to match research aims, or study objectives, to specific methodologies. There is no single, simple answer as to how we bring research to bear in the design process for specifically improving and systematizing user feedback mechanisms, but a variety of techniques is available, which should be incorporated in the intelligent, user-oriented design process.

Online System Monitoring and Transaction Log Analysis

Before concluding this presentation and opening up the topic for discussion, I would like to focus more closely on online system monitoring and explain the methodology and objectives of transaction log analysis.

FIGURE 11

RESEARCH STUDY AIMS

TO DISCOVER:	VARIABLES AFFECTING USER ATTITUDES/BEHAVIOR
	USER ATTITUDES, PERCEPTIONS, PREFERENCES/NEEDS
	ACTUAL USER BEHAVIOR (DISCRETE, SUMMARY, PATTERNS)
	USER SATISFACTION/FRUSTRATION
	USER PERFORMANCE
TO CONFIRM:	HUNCHES, THEORIES, HYPOTHESES, "REQUIREMENTS"
	(E.G., EXAMPLES IN HELP DISPLAYS REDUCE THE USER'S ERROR RATE; SUGGESTIVE PROMPTS LEAD TO MORE EFFICIENT SEARCHING; FORCED BROWSING AMONG HEADINGS IMPROVES RECALL)
TO EVALUATE:	USER OR SYSTEM PERFORMANCE FACTORS
	MODIFICATIONS TO THE DATABASE STRUCTURE, RETRIEVAL SOFTWARE, OR THE USER INTERFACE
TO PREDICT:	INTERACTION BETWEEN A VARIETY OF USER/SYSTEM VARIABLES

FIGURE 12

STUDY METHODOLOGIES

(DATA COLLECTION, DATA ANALYSIS)

* LITERATURE REVIEW
* INTERVIEWS
* SURVEYS/QUESTIONNAIRES
* ANALYSIS OF SECONDARY DATA
* INFORMAL OBSERVATION & CONSULTATION
* SUGGESTION BOX, COMMENTS FACILITY
* PANEL OF EXPERTS
* CONTROLLED, LABORATORY TESTING
* FIELD PERFORMANCE TESTING
* UNOBTRUSIVE MONITORING
* TRANSACTION ANALYSIS

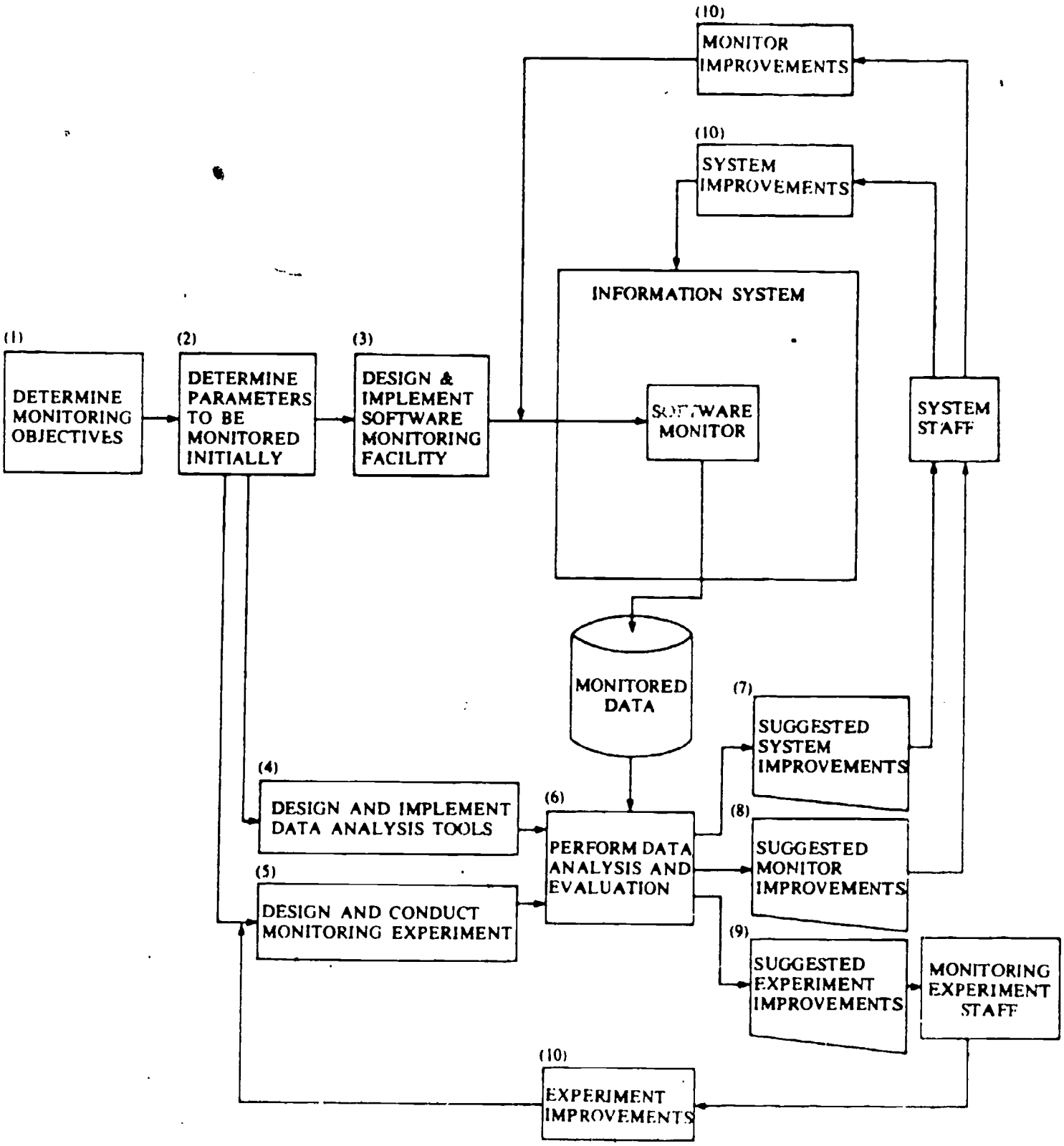
Penniman and Dominick have discussed system monitoring and evaluation techniques in some detail. Two purposes for monitoring computer systems are, generally speaking, to track system performance (e.g., response time, throughput, identify bottlenecks, etc.) and to track and record actual user activity on the system, especially on interactive systems like online information retrieval systems. Special purpose software can be devised to reside within the computer system to carry out these tasks. System or user activity is typically captured or mirrored in some sense, and recorded on a log (some storage medium) by the software monitor. The activity to be monitored, and the resultant data to be written (transferred) to the log, depends upon the objectives identified for the monitoring process. Activity on any computer system consists of many varieties and levels, and, taken in its entirety, can be voluminous and unmanageable. System monitoring has associated processing and storage costs, so selection of types of data to be collected on the log must be taken very seriously. The "phases" or steps involved in the interactive system monitoring and evaluation process cycle are well represented in Penniman and Dominick's schematic (Figure 13). This structured process can be applied to either system activity elements or user activity elements, depending on the research/design objectives and the specific questions to be answered. The ultimate goals of the process are improving system performance and maximizing the utility of the system for its users.

Online system usage monitoring and analysis of user activity data recorded on system logs is especially valuable because it can tell us, with a great deal of accuracy, exactly what users of the system actually do, what actually happens online, in as much detail as we prefer. Once the monitor is developed and plugged in, the amount of data to be collected is limited only by the amount of log storage capacity available. In other words, the data collection sample can be large or small, comprehensive or selective.

Research based on system monitoring of online information retrieval systems, including online library catalogs, has been conducted by a number of investigators.[9-13] The data collected for analysis are now commonly referred to as transaction logs, and a set of methods used to study such logs is known as transaction log analysis. While the objectives of transaction log analysis are commonly understood, unfortunately the data captured and logged vary from monitor to monitor and system to system. Users and systems vary in ways that may be beyond the researcher's control, but a great degree of uniformity can be achieved in the monitoring of user-system activity (i.e., the content and format of system logs). Software monitors have the ability to log every user input action and every system action or response, both apparent and transparent to the user.

The two general objectives of transaction log analysis are to discover how a given system is being used--what actually happens vis-a-vis user-system interaction--and to determine any patterns of use among a given population

FIGURE 13: Monitoring and Evaluation of Online Information System Usage (PENNIMAN AND DOMINICK, 1980)



of system users. John Tolle has identified a number of questions that can be answered from transaction log analysis.[13] These include:

1. What commands were used, and with what frequency?

2. What errors do users make, and where in the online session do they make them?

3. What types of searches are conducted?

4. What are the number of retrievals resulting from various search types?

5. How much time do patrons spend at a terminal?

6. How do times per transaction set vary?

7. How does use vary among terminals, locations, or over time (seasons, semesters, etc.)?

8. What are the patterns or probabilities of going from one context or stage to any other?

9. Do searchers change search types?

10. Which system features are used frequently, or infrequently?

11. What system state (e.g., message response, entry requirement or prompt) precedes frequent user errors?

12. What is the tempo or pace of interaction at various levels or stages of the search process?

And many more! Indeed, the most exciting prospect of conducting transaction log analysis is discovering patterns of online searching (i.e., the probabilities of a user proceeding from a particular current state or command directly to a particular succeeding state or command) and tracking those patterns of use over time, or across user populations and terminal locations. The goals of the approach should be obvious. The assumptions and objectives that underlie system design can be assessed by comparison with actual system use. Costly features favored by designers may not be used. "Easy to use" entry methods may not eliminate user errors. The limitations of good system design, alone, may be discovered. As Cochrane and Markey sum it up, "From studies of transaction logs can come system improvements, better user assistance, possibly better staff and user interfaces...and a better understanding of online catalog use."[14]

Transaction Log Analysis Methodologies

What is required to make transaction log analysis as effective and informative as possible? To answer this, a little more must be understood about the methodology of transaction log analysis.

Certain minimum data collection requirements must be met for transaction log analysis to be useful. This data is captured by a monitor and recorded on a machine-readable transaction log. Since the amount of the data stored in the log is so voluminous, computer programs must be written to read and analyze this data in an efficient manner. During the initial reading of a transaction log (usually stored on magnetic tape), data validity checks, data reformatting, data decoding, or data reduction activities may be performed. Much of the raw interaction data logged from user sessions is irrelevant to the analysis procedures. Even if one wishes to regenerate user sessions in their entirety in print form, a week's worth of online sessions would easily fill a large room with paper. Thus, for machine-based analysis of transaction logs, data encoding and reduction techniques are usually employed. Commands and screen messages, for example, can be coded and stored using only a few bytes of storage space, rather than storing the entire command character string and the full system display or message.

Even if a universal monitor were available, each online system is different at the functional, transaction level. Thus, transaction logs from different systems will be different in content, if not in format. Since investigators can only count and analyze what is recorded on the log, the following data elements should be included to permit analysis that will be useful and informative:

1. Terminal (on which transactions took place) identification
2. Service/database/file searched
3. User command/request
4. User search term/text entered
5. Time of the user entry (to the second)
6. System response
7. Time of system response when made apparent to user
8. Starting time of user session (if available)
9. Ending time of user session (if available)
10. Date of above events

User commands and system responses are usually coded on transaction logs, so frequency counts for each of these events can be obtained fairly easily. These raw counts can answer a number of questions about system use. But online retrieval systems are designed to permit intelligent dialogue between the user and the system. Understanding of the user-system interaction is achieved by viewing certain discrete events in their logical combination. These combinations can be interpreted as purposeful contexts or "states" in the search process. States consist of one or more user or system events, usually in some specific combination with one another. To facilitate higher levels of understanding of user-system interaction, and the analysis to support it, one or more state taxonomies (a logical set of states) are developed by the investigator. A state taxonomy can be short (BEGIN, SEARCH, DISPLAY, END) or long, but usually represents a careful reduction of all events logged, a reduction achieved through a mapping of user and system events to their appropriate states.

Obviously, the development of state taxonomies for online sessions and the correct mapping of events to states, require a solid understanding of the online retrieval process and, specifically, a comprehensive knowledge of a particular system's functionality: commands, dialogue structure, displays, and messages. Since there can be more than one desirable state taxonomy for purposes of analysis, there can be several "correct" mappings of the discrete user/system events to specific states. Incorrect mappings are possible, of course (commands do not always do what their names indicate, and may perform differently in different search contexts in the same system), so an understanding of the system and the value of certain events in specific transaction contexts is essential to the mapping process and the interpretation of the results of higher-order (beyond frequency counts) data analysis. This complex process within transaction log analysis is represented in Figure 14.

One of the online retrieval systems at the Library of Congress (LC) is named SCORPIO. Many commands are available in SCORPIO, and they are listed in the rightmost column. Since online retrieval systems can differ considerably in functionality, a particular system's commands must be interpreted in the light of more general search process steps (leftmost column). An eleven-state taxonomy was created to permit stochastic process (probability) or state transition analyses of user sessions conducted on SCORPIO. Particular commands, many alike in search process function, were then mapped to specific states in the taxonomy.

To capture the flow and pace of online sessions, it is necessary to identify well-bounded user sessions from the transaction logs, and assign state codes to all the relevant events (user or system) recorded on the log. In systems that require the use of begin and end session protocols, defining user sessions is a simple matter. In cases where session boundaries are not well-defined on the transaction log, certain timing algorithms can be used to create a reliable sample of user sessions for analysis. For example, 4-5 minutes between user inputs is a good indication that the first user has left the terminal.

FIGURE 14

Primary Mapping SCORPIO - LC States/Commands

PROCESS	STATE MAP	LC COMMAND
1. File Selection	BGNS	BGNS
2. Term selection/review No sets created, this is an index review No display of records operates only within an index or thesaurus (by definition)	BRWS	BRWS LIVT
3. Create sets (simple search)	SRCH	SLCT EXPN
4. Reformulate, negotiate, refine query (combine, limit, etc.) where you have actually done a search of the database	MODS	COMB LIMT
5. Display records/citations	DISP	DSPL NEXT "XMIT"
6. Review search history/search status obtaining information about the history and/or status of the session we are in View system/session information cost show settings etc.	STAT	HIST RLSE SHOW NEWS SET
7. A help function with its own command	HELP	No Command
8. End session cost might also be here	ENDS	ENDS
9. COMBINED FUNCTIONS: FIND = BRWS/SLCT/(COMB)/DSPL SCAN = SLCT/DSPL	FIND	FIND SCAN
10. COMBINED FUNCTIONS: RETR = SLCT/COMB	RETR	RETR
11. ERROR	ERR	Not a valid command

Once sessions are identified and online events are mapped to various states, a variety of state transition time-based or probability analyses can be conducted. Figure 15 represents the findings of a state transition analysis of a sample of SCORPIO user sessions. This state transition matrix shows, in percentages, the frequencies of transitions from a "current" state (rows) to an immediate successor state (columns). Figure 15 shows, for example, that in 59.8% of the cases recorded and analyzed, the user proceeds from one error state (ERRS) to another error state. Yet relatively few errors are made (2.5%) in the display (DISP) state.

Different state taxonomies--basic, extended, refined--used for transition probability analysis permit a variety of levels of understanding of system use. A crude taxonomy (e.g., SEARCH, DISPLAY) limits any useful understanding of system use, or cross-system comparisons. On the other hand, an extended taxonomy can result in incomprehensible findings (imagine a transition matrix with 50 rows and columns). Simple, inflexible online retrieval systems may incorporate a small set of meaningful transaction states. Analysis of such systems and their use is relatively easy, but findings may offer little enlightenment. However, even in systems where only one short search path is permitted, state transition analysis can tell us, for example, how frequently users complete the process, where and with what frequency they "bail out," and where they are most likely to make errors.

Transaction log analysis has limitations. It will not yield answers to questions about user intentions, attitudes, or preferences. It cannot help us determine the quality of specific search results (e.g., recall or precision). Cross-system comparisons are difficult because the structure of the search process--what is permitted, what is guided or encouraged--may differ considerably from system to system. State taxonomies can be manipulated to facilitate comparisons across systems, but caution must be exercised in the interpretation of such findings. Important, determining variables, subtle in online search environments, may be overlooked or buried within coarse transaction states.

Properly conducted to achieve relevant research objectives, transaction log analysis can be one of the more effective means of providing valuable feedback within the total online system design process.

FIGURE 15

Transition Probability Matrix
Primary Mapping - 1st Order - Context Validated - Successor - LC
March - June 1982 (Expressed in Percent)

	BGNS	BRWS	SRCH	DISP	FIND	RETR	MODS	STAT	ENDS	ERRS
BGNS	15.5	64.3	0.0	1.1	1.9	0.1	0.1	0.3	9.4	6.7
BRWS	1.0	48.3	28.3	4.4	8.9	0.3	0.1	0.6	4.8	3.4
SRCH	0.4	6.5	15.6	67.7	0.9	0.1	0.9	1.2	1.7	5.1
DISP	0.3	9.5	3.1	77.4	1.0	0.1	0.2	1.5	4.4	2.5
FIND	0.3	11.0	1.8	5.5	67.9	0.0	1.4	1.0	3.7	2.5
RETR	0.0	33.3	5.6	38.9	0.0	16.7	0.0	5.6	0.0	0.0
MODS	0.0	8.2	0.0	30.9	3.3	0.0	54.5	2.4	0.8	0.0
STAT	1.3	4.7	1.3	56.4	2.0	0.0	10.7	10.1	11.4	2.0
ENDS	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
ERRS	2.4	10.6	3.4	13.4	0.5	0.0	0.2	1.0	8.7	59.8

(Tolle, 1983)

References

1. Ray R. Larson and Vicki Graham, "Monitoring and Evaluating MELVYL," Information Technology and Libraries 2 (March 1983): 93-104.

2. W. D. Penniman and W. D. Dominick, "Monitoring and Evaluation of Online Information System Usage," Information Processing & Management 16 (1980): 17-35.

3. Milton J. Marcus, "Concerning the Nature of Information Processing End Use," in The Role of Human Factors in Computers, ed. Richard E. Granada and Jay M. Finkelman (New York: Human Factors Society, New York Metropolitan Chapter, 1977).

4. Charles Hildreth, Online Public Access Catalogs: The User Perspective (Dublin, Ohio: OCLC, 1982).

5. Pauline A. Cochrane, "Where Do We Go From Here?" Online 5 (July 1981): 30-41.

6. Penniman and Dominick, "Monitoring and Evaluation," 18.

7. Ben Shneiderman, Software Psychology: Human Factors in Computer and Information Systems (Cambridge, Mass.: Winthrop Publishers, 1980), 258-59.

8. Pauline A. Cochrane and Karen Markey, "Catalog Use Studies -- Since the Introduction of Online Interactive Catalogs: Impact on Design for Subject Access," Library and Information Science Research 5 (1983): 337-63.

9. Janet L. Chapman, "A State Transition Analysis of Online Information-Seeking Behavior," Journal of the American Society for Information Science 31 (September 1981): 325-33.

10. Michael D. Cooper, "Usage Patterns of an Online Search System," Journal of the American Society for Information Science 34 (September 1983): 343-49.

11. Thomas H. Martin, John C. Wyman, and Kumud Madhok, Feedback and Exploratory Mechanisms For Assisting Library Staff Improve Online Catalog Searching, Final Report to the Council on Library Resources (Syracuse, N.Y.: Syracuse University, 1983).

12. W. David Penniman, Modeling and Evaluation of On-Line User Behavior, Final Report to the National Library of Medicine. Extramural program grant no. NLM/EMP(1 RO1 LM 03444-01) (Dublin, Ohio: OCLC, 1981).

13. John E. Tolle, Current Utilization of Online Catalogs: Transaction Log Analysis, Volume 1, Final Report to the Council on Library Resources. OCLC Office of Research. Report No. OCLC/OPR/RR-83/2 (Dublin, Ohio: OCLC, 1983).

14. Cochrane and Markey, "Catalog Use Studies."

QUESTIONS AND DISCUSSION

Comment: A couple of important things are missing from your five stages of the design process. The database itself should have been included along with the system knowledge and training and user requirements. I think we must pay much more attention to the data itself, because there are a lot of things implicit in the data that never have been taken advantage of. It's only these days that more attention is being paid to that. Of course, the database management system people always paid more attention to what is inherent in the data.

Second, as a driving force, technology itself is forcing system designers to make all kinds of changes.

The third thing is external systems. You are often forced, whether you like it or not, to accommodate different protocols, different sources of information, to tap into things external to your own immediate system. That is also an important component here.

Hildreth: Sure. I had a very narrow purpose for the series of five slides. Certainly there's a lot that could be added to the boxes on the design process diagrams. A lot of things can be separated out. I don't think this is complete, but it's helpful for expository purposes, to bring out more of this.

Question: It sounds like most of the data needed for transaction analysis is readily available in the system, and I assume we're writing this stuff off onto a tape drive. Are there any economic or other reasons why, in a given system, someone wouldn't log some of these things, or is it pure over.ight?

Hildreth: Pure oversight. (laughter) Sometimes, of course, it's overhead. There's a cost in doing this. There are answers to that problem. Don't do it all the time; do it once in a while. There are different approaches to it. But sometimes it is oversight.

Question: Are you suggesting that users do this analysis, rather than the people who designed the system?

Hildreth: The users don't do any part of this analysis. They're done. Of course, if you don't understand the system processes from the user's point of view, that is, use it a lot yourself, your analysis probably won't be good. In order to get our RAs to duly grasp the systems, we make them go over and dial-up and use one of these systems quite a bit. Granted, all the insights came from using the systems themselves.

Question: When you say RAs, you're talking about your research assistants?

Hildreth: Yes.

Question: Aren't they, by definition, users of the system? They aren't the people who designed them.

Hildreth: Right.

Comment: Okay. So then the answer is yes; users did this aralysis, not the system designers.

Hildreth: Oh, now I understand what you mean by that. Yes. I've seen the results of this kind of analysis done by non-users of the system. It's usually incorrect.

There's not a single correct mapping. I think that should be intuitively clear. It depends on how refined you want your analysis to be. Is this the correct mapping or taxonomy? Should I reduce those by coming up with more general categories, collapse them into three or four states, and then do the remapping? It depends on how much you want to know, and what you want to know. So I don't think there's any single correct mapping. But there are incorrect mappings, of course. The first time this analysis was done, there were incorrect mappings, because they misunderstood what certain of the LC commands did in context of actual use. If you look at the semantics of a command, the written documentation, how it operates within a system, then you're likely to do the mapping correctly.

Comment: But there is only one correct mapping, because computers are Turing machines. If you look at the source code--

Hildreth: More than one correct mapping.

Comment: No, no. Look at the source code. Given that the computer will follow the same set of instructions, there is only one. It may be very very difficult to get it, but the point is that there does exist, for version X of the system on this day, a correct mapping.

Hildreth: Let's call it a basic mapping, because correct is relative to what you want to get out of the analysis. I understand what you mean. But since there may be so many activity codes or state codes that you're dealing with, the basic mapping is too cumbersome for most systems, so we have to do some collapsing into one or more of what I call correct mappings. Correct relative to the objectives of the analysis, and what you want out of it.

Question: Would it not be important to consider building into systems facilities that allow the user to provide important feedback such as, if the user looks at an item, is it important to the user or not--the Salton type of relevancy? Or, if you started with a set of keywords, and if your system does suggest subject headings, then when you allow the user to choose from them,

why not give him the option of indicating what are the most important ones, either implicitly or explicitly, by order of selection numbers or other mechanisms? Would this, do you think, fit into the feedback interpretation?

Hildreth: Yes. That was the original bifurcation of feedback mechanisms. I knew I was going to go one way or the other in the presentation. The other interpretation was the online, real-time, interactive kinds of feedback mechanisms you've just referred to. Ironically, and this is probably intuitively clear to all of you, the bifurcation, whether one follows one path or the other, ultimately comes back together. Let me explain. As we refine some of these analysis techniques for logging, some of the actual analysis programs themselves are going to be built back into the real-time interaction, just what you're talking about, to provide the basis for some intelligent guiding, or what Penniman called adaptive prompting. When Penniman published his report on the ELHILL study, his whole point was that this will be useful, not so much for ex post facto analysis, but for building in later into the real-time systems and interfaces.

Comment: I think that the end user, the public that does the searching in the catalog, brings something special to the search process that the trained intermediary simply cannot provide. Despite what some authors say, that user does know at the moment, in the context of his or her search, what makes sense, what's important for them, whether it's a record or a term or so on. Unless we capitalize on this kind of feedback, we are missing out on something very special that only that user can give.

Hildreth: I think the point that Penniman made is that if we're to have much intelligent guidance in what we call adaptive prompting, in real-time, using individualized instruction procedures such as you experimented with in the IIDA project at Drexel, you have to have diagnostic procedures and analysis routines built in. They have to go much further than we do now, which is really pretty much context-independent checking for semantic and syntax errors. A requirement for doing further kinds of adaptive prompting and intelligent guidance is that the system understand what the user is doing. So we have this overlap there, research and actual real-time intelligent assistance. But to do that, it's got to know the historical context, where the user is at that point. That's why some of these insights into state transition and process analysis, as we refine them and discard what's not so useful, come back into the design of real-time processing.

Comment: I think that idea is really valuable for two reasons. First of all, it's useful for the identification of what we call the working set. Secondly, in doing transaction analysis and developing understanding of state transitions, you want to know which transitions turn out to be the most successful. If you don't know that, then you can't provide easier ways to get through those transitions, to direct people toward the more successful types of interaction. So I think they go hand in hand.

In the process of data capture, we have to pay some attention to being able to define the scope of the logging activity at any one point in time. Maybe, for example, we should log only certain functional areas of the system. We may wish to randomly sample sessions across the whole system. We may wish to restrict the sampling to a certain population at any one point in time. Just turning on a log and capturing everything is often not very productive.

Hildreth: Everything can mean, of course, two things. Everything all of the time, or everything input and output for a given session, but not every session. Of course, there are options for both of these.

Comment: The other thing in the way of data to include is resource consumption. That could provide an awful lot of very useful information. If you're logging all the things that you mentioned, throwing resource consumption in should be easy.

Comment: Charles, I can appreciate what you're proposing here. I still have some questions as to whether it's going to yield just what we hope it would. Is it going to tell me what I need to know about the end user? It's an attractive route to pursue, because it's more convenient to collect that data if we use the stuff at hand. We have the hardware technology to get that. So we measure what we can measure.

My problem with it is that, as we refine it and agree on the questions, I'm still not sure that they'll be precise enough to tell us the kinds of things we need to know about the value of search options. You can tell me that 59% of the time that somebody made an error in Scorpio, they went to another error state. I can change my error messages and retest the frequency distribution. But I have some difficulties accounting for changes in population. In the time it took to change the error messages, that population is also changing.

I can look at specific kinds of things. But to tell me that we've got to go from one browse level to another browse level--without the context of understanding what the user was doing and why, I really am hesitant to draw conclusions from it, to say that I am understanding the search functions.

I feel like I'm doing it through three or four mirrors. By the time the reflection gets to me, it may be interesting data, but I'm not sure what I can do with it.

Hildreth: Your phrase, understanding the search process, is key. No single approach is going to give us that. I can think of other approaches that helped me understand the search process much more--protocol analysis, actual interviews with the users, controlled experiments. That's why I deliberately couched this last bit on transaction analysis so I wouldn't be accused of putting too much in it. I wanted to funnel down to call for some standards, at least some uniformity.

There are the other methods and they have their value. I certainly don't mean to indicate that state transition analysis is going to give me that feel for how the system is being used. It doesn't answer most of the why questions. Frequencies tell us something about whether a feature is being used. But, as you rightly point out, any time you're designing retrieval software and interface software, you've got a definite logical structure. You make some heavy logical and transition type decisions. This analysis can tell us if your assumptions were correct, whether it leads to where you thought it would, whether it's going to be used in the transitional process the way you thought it would so it's very useful information to the designer, but in a very narrow way.

Question: You introduced something called user trainers at one point. Do those people really represent what the users want? Shall we pay much attention to them, or are they misleading us?

Hildreth: Yes and no. We have to pay attention to them.

Comment: Yes, we sort of work for them. (laughter)

Hildreth: The question is complex, obviously. I think if we don't listen to the reference staff that's interacting with the public, especially at the time of the introduction of the catalog, we're missing a lot of important information. A lot of that same information can be gotten by controlled experiments, the things we're doing now at OCLC, where we're manipulating some interface variables, and sitting down with users and having them go through search tasks. We're recording and listening to the verbal protocols, and analyzing it later. But reference librarians, when that new online catalog is put out there, are hearing that daily too.

Question: But are they telling us what the users want? Or what they think the users want?

Hildreth: Well, it's complex. They certainly filter the user-responses.

Comment: Well, maybe we should just ignore them, and get it directly from the users.

Question: Another way to state the question is, are the reference librarians as good at intuiting what the users want as the systems designers sitting back in their offices? (laughter)

Comment: Why not make the assumption that they are? Then we have additional bits of information to guide us. If we make the assumption that they're not, we're just closing ourselves off.

Hildreth: There also has to be a checks and balance process in there, so that we don't go too far with the analysis of that data from one particular method or feedback channel. Some system designers, because they like the reference librarian, for whatever reason, may pay too much attention to what she's saying about user reactions.

Comment: One of the other advantages of standards would be that they would help some individuals in interpreting their data. Too often, they ask for data and have no idea how in the world they would ever use it after they got it.

But I wonder if, in the logging process, we shouldn't merely ask the person at the end whether or not they got what they want, but rather, during the process, ask if they are finding or have found what they want. Very often, people will continue browsing even after they have found the very thing that they want, for some reason--just because it's fun, maybe because they found what they wanted and then realized that there's a heck of a lot more there than they thought at first. So I would think that you're not merely going to ask for some feedback from them at the end of the process, but also during the process--and also record that.

Hildreth: With controlled experiments, that's fairly easy to do, and you get their permission. In real usage, if you intrude and pop up a question or something for feedback while they're searching, you're into sticky problems of privacy and intruding on their thought processes. I think it's a great idea. How to administrate it is a difficult question.

Comment: I think one of the problems we've got with online catalogs, and I'm making a distinction here between that and something like DIALOG, is that the users aren't the customers. You see that most clearly when you're trying to sell something. The problem is that the people who use the online catalog are not the people who specify it. They're not the library planners, and they're not the people who are going to decide if it's a success.

Oh, there'll be some feedback from them. The faculty member, who's on the library committee, who comes in and yells at the chief librarian, will have an impact. But he may not represent the 2,000 undergraduates who like that feature that he's upset about.

So part of the difficulty in terms of user feedback, and I see this in a commercial environment but it exists at all levels, is that the people that we're trying to serve with the online catalog are not the people who are defining the success of what we're doing and controlling that process. There are a lot of features that we're putting into our online catalog that we think are useless--except for one thing. There're going to bring in dollars, because the people who are making the decisions say they're necessary. They may be right or wrong.

Comment: You are saying that the users are not the people who pay for it.

Comment: Absolutely. That's different from something like DIALOG because there the users are the people who pay.

Comment: Well, in an academic environment the students ultimately pay for it.

Comment: But only indirectly.

Comment: In a meeting that I was at on Wednesday morning, a reference librarian made some rather derogatory comments about systems designers. They're told what the librarians want, they rush off to their ivory tower, they come back and deliver something that doesn't look anything like it was supposed to look like. I think that's still going on. I'd like to interject into the design process a prototyping mechanism that has a lot of iterations early on, instead of waiting until we get to the end and then doing all of the analysis of use of the system. We should get feedback real early in the design process, so that the end result looks a little bit more like what the user wants by the time we get there.

Comment: Yes. There is a fundamental problem with that, however. As you provide prototypes to people, no matter how many times you explain to them that it's a prototype, basic underlying feelings about the way that system operates are formed from that initial exposure. A lot of the people that you ask for analysis from will be disappointed just by looking at the prototype, and will fail to give the input that you want. It requires understanding of the possibilities.

Comment: Not if the versions come quickly enough after their negative reactions. There can't be five months between. I'm talking about two days or three days between prototypes.

Comment: You need some high-level development tools to do that.

Comment: Yes. Those are very important. You should be able to zip up screens in a matter of minutes, and make those kinds of cosmetic changes. That is really what the user sees; he only has a perception of what is displayed on the screen.

Comment: These tools are increasingly available, and they do make a very big difference. You can sit down, whether you're a librarian or whoever, and mark up screens on the fly and try things out in a matter of minutes.

Comment: I ran across a very interesting article this week, which I ripped out. The title was, "Librarians: The Untapped Resource." It came out of the September Datamation. It was basically saying that a lot of computer people trying to establish databases do not understand the process of drawing

out the user to determine exactly what it is they're after. That's a very well-known phenomenon among reference librarians. It's almost as though the user is hiding what it is that he's really wanting. What is there, in the transaction log analysis, that might suggest that we're any better at pulling out exactly what it was the user was after than an interaction or dialogue between the reference librarian and the end user?

Hildreth: We're throwing around this phrase "transaction log analysis" a lot now, especially at OCLC, and people may think that there's one technique. There are a number of ways you can analyze transaction logs. Machine analysis, stochastic process analysis isn't going to help us a lot with what you've just been talking about. You get a feel for it, some incredible insights when you eyeball hard copy or run it back up on the screen, and walk right through what the users have done. I think that's very important, really useful, and one just has to keep it in perspective.

Comment: It's like a lot of other things we've talked about. You just can't swing on one branch. (laughter)

Comment: Especially if you're falling on your own sword. (laughter)

Comment: That kind of analysis will really pose more questions than it answers. You've got to know which questions to ask when you go out to interview or look at a transaction log. You can't just go fishing.

Comment: That's why I'm sorry to see intuition disappearing in your charts. We've wanted to get users, both library staff and real users, involved in the design process, and it's been very, very difficult. For the most part, they haven't the foggiest notion of what they need or what they want.

Comment: That's why prototyping does it, because they can't conceptualize very well until they have something to beat on.

Comment: And even when they have something to beat on, they can't always say, "Well, this is great, if it could only do X." Sometimes they can, but it's very difficult to get them to envision something that doesn't exist.

Comment: You've got to push enough of them past it, so that you get the ones that can carry out that thought process. Not all of them can do it.

Comment: One thing that's important to remember is that the vast majority of users never see a reference librarian. We've never been able to touch them. It's a very special set of users. I don't know what the characteristics of the set are, but it's a very special set of people that ever talk to a reference librarian. That's got to be skewed. I don't know how, and I don't know whether anybody else does, but I'm positive that it is.

Comment: I just want to follow up on getting some kind of a log, a straight log of what people have done. I forget who told me that they had done this, but someone rigged a slave printer to a terminal, hid the slave printer in another room so the user didn't know what was going on, and ended up with a huge, voluminous printout of what happened. I would suggest turning that over to a reference librarian. It might answer some questions about what is going on, what approaches are being taken. I believe there is a CLR grant that Brian Nielsen has at Northwestern to do some work in this area.

Comment: I'd just like to pay tribute to what you're doing, and remind everybody that this kind of work is very humbling. What we're basically dealing with here is human behavior. Humans are marvelously diverse creatures. There must be an uncertainty principle pertaining to this, because I think you'll have to start out by accepting that you cannot do the job that you set out to do.

You want a lot of information, you want to sample during a search, fine. That is great and useful, but it's also intrusive. If you want controlled experiments, fine. But then you really don't know how these people will behave in the real world. You want a nice, uniform body of people, fine. Grab Library Science 103 some afternoon. Then you've got a totally artificial situation--but the standard deviation is nice and small. (laughter)

The point is that you really can't get there from here. But it doesn't mean that you shouldn't keep trying. We just must accept that there never will be a model that is mechanically predictive of total human behavior. We just can't handle it. So I think the work you've done is fine. Keep going. I don't mean to be negative. We must keep going. But accept that this is not like a mathematical theorem-solving problem. It will not come to the end.

V. SCREEN LAYOUTS AND DISPLAYS

Joseph R. Matthews

The display of bibliographic information on a CRT screen raises two basic issues: first, the format and arrangement of specific data elements on the screen, and second, the amount of information that is to be displayed for a particular user. The screen layout or display of information on a CRT screen constitutes the most important ingredient of a "user-friendly" interface; yet this is perhaps the most frequently overlooked element of system design.

There are a few guides for screen layouts and displays providing suggestions to designers of online systems.[1] Little of this advice is based on systematic controlled research. Often, designers of online catalogs assume that the arrangement of information on the screen is a relatively straightforward task. Decisions in this area are dictated by tradition, i.e., the typical card catalog arrangement is adopted for display of bibliographic information, or the task is left in the hands of a computer programmer. It is my belief that, since the user of an automated system spends the majority of time reading information displayed on a CRT screen, careful consideration of options for screen layouts and displays should assume a much more important role in the design of online catalogs.[2]

The traditional, standardized 3x5 card found in most library card catalogs has been with libraries for so long that knowledge of the evolution of that particular medium has left the consciousness of the profession. It is also assumed by librarians that the display of bibliographic information is understood and appreciated by the user of the card catalog. The layout, punctuation, arrangement of data elements, and spacing of data elements on catalog cards, while broadly uniform from catalog to catalog, is nevertheless more a mystery to the average user than the profession would care to admit. Only 59% of library users in all types of libraries use the card catalog. Known and unknown item searches occur about equally, and subject searching occurs more frequently in public libraries than in academic libraries.[3]

Alternatives to the card catalog, i.e., the book catalog, the COM catalog, and most recently the online catalog, offer the opportunity for librarians and designers of online catalogs to develop systems that display bibliographic information according to the needs of users rather than the assumptions and traditions of the library profession.

The typical card catalog format, as seen in Figure 1, is most familiar to those of us who have worked with libraries for a period of time. Yet, as

seen in Figures 2-7, there is a variety of ways in which bibliographic information has been displayed via online catalogs.

Some posit that the display of bibliographic information on a CRT screen is not yet thoroughly understood and should be given more attention. Jim Dwyer has suggested that both COM and online catalogs seem to be tied in part to the format and jargon that have worked so poorly in card catalogs.[4] The "standard" library record format, including indentations, bold or capital letters, red subject headings, abbreviations, symbols, and punctuation can suggest to the uninitiated that they are confronting a foreign language--professional "librarianese." Neville has characterized this special language as "biblish."[5] It is, for all intents and purposes, unintelligible to the average user. Why then do we persist in using it?

For a public access online catalog to be a tool that truly meets the needs of the public, the needs of the user must dictate the format and amount of information that is displayed.[6] To be truly effective, the visual presentation of bibliographic information should grow out of the logical structure of the data and facilitate its use by patrons[7]--in other words, the design of online catalog screen layouts and displays should be user driven.

Users surveyed in the recent CLR-sponsored Online Catalog Study indicated that the majority of the interface problems that they experienced were concentrated in the areas of search formulation and understanding and control of displays.[8] They had trouble making the catalog tell them what they wanted to know. This should tell us something!

One of the reasons users cannot get catalogs to work is that only a portion of the information displayed in each record is relevant to their search requirements. Sometimes the data sought by the user are buried in the middle of a record. Online catalog technology offers many opportunities--as yet unexploited--to tailor search results to match users' needs.

When designing an online catalog and deciding the amount of information that is to be displayed, it must be remembered that all online catalog users do not bring similar needs to the catalog. Thus, defining a variety of user classifications is both possible and desirable. The most popular approach groups users into three basic categories,[9] for which short, medium, and full MARC bibliographic records are provided (and for the latter, perhaps a display with and without MARC tags and delimiters). It should, however, be noted that terminology describing the possible display options varies considerably. For example, Hildreth notes that one similar group of formats has been variously called: "review, multiple, title, index display, and truncated entry." And in another group, the terms include: "long, full, expanded, total, base, and MARC."[10] Similarly, once a number of bibliographic records have been retrieved, a user can move "up or down," "back or forward," or view the "previous screen or next screen." Why can't we limit this proliferation of catalog jargon?

FIGURE 1: Typical card catalog format

HQ
766
.P636

Pohlman, Edward, 1933-
Incentives and compensations in birth planning. [Chapel Hill] Carolina Population Center, 1971.
vii, 140 p. 23 cm. (Carolina Population Center. Monograph 11)
Bibliography: p. 127-137.

1. Birth control. I. Title. II. Series.

HQ766.P6368 362.8/2
75-635590

EMXTG B/NA A B2-256104 SAMP0039 08/17/83

FIGURE 2: Mankato State

Screen 001 of 001 Record 0002 of 0002 MSU
LOCTN: ERC F229 .C6x
TITLE: Cooking in colonial days: a Williamsberg Kitchen (Filmstrip)
PUBLR: Colonial Williamsburg, 1956.
DESCR: 49 fr., color. 35 mm.
NOTES: Describes the physical location and arrangement of the kitchen and gardens of a typical upper-class home in colonial Virginia. Follows the servants' preparation of breakfast and dinner, showing various types of foodstuffs, cooking utensils, gadgets, and kitchen equipment. Captioned photographs.
SUBJT: Kitchen utensils.
SUBJT: Cookery--History.
SUBJT: Cookery, American--Virginia.
SUBJT: United States--History--Colonial period, ca. 1600-1775.

FIGURE 3: NLM Integrated Library System

SUBJECT: LIBRARIES, UNIVERSITY AND COLLEGE
FOUND: 3

REF	DATE	TITLES	AUTHOR	CALL NUMBER
R1	1977	Developing an acquisitions system	Hindle, Antho	Z689.H55
R2	1977	Approval plans and academic libra	McCullough, K	Z689.M15
R3	1977	Economics of serials management :	Backwell's P	Z692.S5.P7x

(END)

CHOICE: R2

Z689.H15

McCullough, Kathleen.
Approval plans and academic libraries : an interpretive survey / by Kathlee
n McCullough, Edwin D. Posey, and Doyle C. Pickett.
Phoenix, AZ : Oryx Press, c1977.
x, 154 p. ; 24 cm.

Libraries, University and college - United States - Acquisitions.
Library surveys - United States.
Book Selection
Library surveys - United States.
Posey, Edwin D.,
Pickett, Doyle C.,

CIRC
STATUS:

COPY #: 1 AVAILABLE

Enter /AU for author, /TI for title, /SU for subject, /TM for term search.

CHOICE:

FIGURE 4: Ohio State LCS

```
 FBL/1
HF5736K87
KUTTNER, MONROE S., 1929
   MANAGING THE PAPERWORK PIPELINE : ACHIEVING COST-EFFECTIVE PAPERWORK AND
INFORMATION PROCESSING / MONROE S. KUTTNER. NEW YORK : WILEY, C1978.
XVI,
244 p. ; 24 CM.
   INCLUDES INDEX. "A RONALD PRESS PUBLICATION." BIBIOGRAPHY: P. 237-240
SUB: 1. BUSINESS RECORDS  2. OFFICE MANAGEMENT  3. COMMUNICATION IN
MANAGEMENT
LC CARD  :77-15041  TITLE  :2222111  OCLC  :3327400     &9Q780620
PAGE 1 END
```

FIGURE 5: University of California MELVYL

```
Type the number you want below or type HELP,  then press RETURN.
  1. See a LONG display of this record.          3. Begin new subject search.
  2. Begin new author/title search.              4. End the session.
>1

Your search for: subject words ENERGY CONSERVATION HOUSES
      retrieved: 1 book from UC libraries.

1.
Author:      Benson, Verel W.
Title:       A guide to energy savings for the poultry producer / (prepared
               by Verel W. Benson). (Washington) : U.S. Dept. of
               Agriculture, (1977)
             ii, 46, (1) p. ; 26 cm.

Notes:       Cover title.
             Bibliography: p. 45-(47)

Subjects:    Poultry houses and equipment -- Energy conservation
             Energy conservation -- United States.
                                        (Record 1 continues on the next screen.)
```

```
Your search for: subject words ENERGY CONSERVATION HOUSES
      retrieved: 1 book from UC libraries.

1. (continued)
Other entries: United States. Dept. of Agriculture.

Call numbers: UCB  Agricul  SF486.B4  (CU-AGRI)

Type the number you want below or type HELP, then press RETURN.
  1 See previous screen of this display          3 Begin new subject search.
  2. Begin new author/title search.              4 End the session.
>4
```

FIGURE 6: Claremont Colleges TLS

WW
HYMAN BLUMBERG SYMPOSIUM ON RESEARCH IN EARLY CHILDHOOD EDUCATION/WOMEN AND THE MATHEMATICAL MYSTIQUE : PROCEEDINGS OF THE EIGHTH ANNUAL HYMAN BLUMBERG SYMPOSIUM ON RESEARCH IN EARLY CHILDHOOD EDUCATION / EDITED BY LYNN H. FOX, LINDA BROOY, AND DIANNE TOBIN. /BALTIMORE . JOHNS HOPKINS. UNIVERSITY PRESS, C1980. /VIII, 211 p. ; ILL. ; 24 CM. / (STUDIES OF INTELLECTUAL PRECOCITY ; 8)/# WOMEN IN MATHEMATICS - CONGRESSES. /# WOMEN MATHEMATICIANS - CONGRESSES. /# SEX DIFFERENCES IN EDUCATION - CONGRESSES. /# MATHEMATICS - STUDY AND TEACHING - CONGRESSES. /& FOX, LYNN H., 1944 - CN /& BRODY, LINDA. CN /& TOBIN, DIANNE. CN /& AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE. CN /(111)/MAN BLUMBERG SYMPOSIUM ON RESEARCH IN EARLY CHILDHOOD EDUCATION, 8 TH, JOHNS HOPKINS UNIVERSITY, 1976.

FIGURE 7: Data Research Associates

Public Access Catalog Full MARC Display Mon 07/19/1982

AAA-0337 Entered: 06/22/1982 Last Modified: 06/22/1982

Type: a Bibl: m Gvt: Lang: eng Src: Ill: Rep: Enc: 8 Cnf: 0 Ctry:
ToD: s MED: Ind: 1 Mod: Fst: 0 Cont: b Int 1: Fic: 0 Bio:
Dates: 1976

010; ; a 76005123 $ o 02867880 $
040; ; a DLC $ c DLC $ d VHB $
020; ; a 0060115017 : $ c $10.95 $
050; 0 ; a HQ772 $ b .G38 1977 $
082; ; a 155.4 $
049; ; a RJB#VXB9,VBB),UJB*,RABA,0QB9,VXBS,VXB=,VXB1,VBBW,VBB%,UJBJ,UJB7,
UJB+UKBU,UKB3,UKB(,RKBY,RJBU,RABQ,RAB*,PJBE,0QBY $
100; 10; a Gesell, Arnold Lucius, $ d 1880-196l. $
245; 14; a The child from five to ten / $ c by Arnold Gesell, Louise Bates
Ames, and Frances L. Ils, in collaboration with Glenna E. Bullis,$
250; ; a Rev. ed. $

Press orange M for more Marc display, yellow Z for title list

Public Access Catalog Full MARC Display Mon 07/19/1982

AAA-0337 Entered: 06/22/1982 Last Modified: 06/22/1982

Type: a Bibl: m Gvt: Lang: eng Src: Ill: Rep: Enc: 8 Cnf: 0 Ctry:
ToD: s MED: Ind: 1 Mod: Fst: 0 Cont: b Int 1: Fic: 0 Bio:
Dates: 1976

010; ; a 76005123 $ o 02867880 $
040; ; a DLC $ c DLC $ d VHB $
020; ; a 0060115017 : $ c $10.95 $
050; 0 ; a HQ772 $ b .G38 1977 $
082; ; a 155.4 $
049; ; a RJB#VXB9,VBB),UJB*,RABA,0QB9,VXBS,VXB=,VXB],VBBW,VBB%,UJBJ,UJB7,
; ; UJB+UKBU,UKB3,UKB(,RKBY,RJBU,RABQ,RAB*,PJBE,0QBY $
100; 10; a Gesell, Arnold Lucius, $ d 1880-1961. $
245; 14; a The child from five to ten / $ c by Arnold Gesell, Louise Bates
Ames, and Frances L. Ils, in collaboration with Glenna E. Bullis,$
250; ; a Rev. ed. $

Press orange M for more Marc display, yellow Z for title list

Given the lack of communication to date among designers of online catalogs, a standardized terminology with which to describe various features, displays, and capabilities of online catalogs is not likely to be achieved in the near future. Such a standard seems like such a simple way to make these new tools easier to use and understand. It's about time system designers quit reinventing the wheel and agreed to cooperate, at least at this minimal level.

Improvement can be made in a variety of areas. While the following discussion will focus on bibliographic records, the advice and suggestions are equally applicable for the display of holdings and status information. There are seven aspects pertaining to CRT display and layouts that must be addressed. These include:

- o layout.
- o content and sequence.
- o vocabulary.
- o typography.
- o spacing.
- o punctuation.
- o color.[11]

Each will be considered in turn.

Layout

The layout--the manner in which information is formatted or located on a CRT screen--is crucial to the effectiveness of an online catalog. Two recent studies highlight the importance of layout. Both studies examined users' preferences in formats for the display of information. Tullis, in an experiment with four different CRT display formats, measured the speed and accuracy of users' interpretation of the information displayed on the screen.[12] The formats included a narrative form using ordinary English, a structured tabular format, and two that used color graphics to differentiate and highlight the data elements. The format in which the information was presented had a clear effect on both the users' performance and their preference for a particular format. Tullis found that both the structured format and the graphic formats resulted in significantly faster user response times than the narrative format. However, Tullis recommended the structured text format as providing the best combination of human performance benefits and lower cost of implementation. For the design of alphanumeric displays, he recommended that:

- o Key information should be presented in a prominent location.
- o Logically related data should be clearly "chunked" and separated from other categories of data.
- o Information should be presented in a fixed, tabular format such that users will develop special expectancies.
- o Presentation of information should be concise.

In a more recent study conducted in a library setting, Benjamin Fryser examined patron preference for versions of a single CRT display bibliographic record: 1) the standard Library of Congress format, or 2) either a table of contents layout or a vertically arranged and underlined field label presentation.[13] These formats were displayed in a random manner to the subject, one of the experimental formats always being paired with the traditional card catalog format. Fryser concluded that for the 347 students studied, both the table of contents and the "labeled, underlined" formats were significantly preferred over the standard Library of Congress card catalog format. Thus it would appear that imposing the traditional card catalog format onto a new technology, i.e., CRT terminal screens, needs to be examined carefully; from the users' perspective it is not the preferred choice. In the light of such findings, why do we persist in using it?

Consistency in screen displays is also helpful since it can reduce user confusion and frustration by allowing skills learned in one situation to be transferred to other situations. In some online catalogs, the screens frequently do not look like one another, usually because the computer handles different searches in different files in different sequences. Results are displayed according to what the computer has found first. This variation in screen displays means that users are forced to figure out the format of each new display before they can interpret the information content of the response. Consistency is not difficult to achieve. When will we insist that the requirements of users take precedence over machine processing requirements?

Consistency is the foundation of a system that is easy to learn, easy to use, and easy to remember. The consistency of commands, system behavior, and display screens allows the user to form a simple and easy-to-remember conceptual model of the online catalog. When the conceptual model of the system from the user's perspective closely parallels that of the designers of the online catalog, then the system can be called user friendly.

Users often want to locate the call number, appropriate librarianese when the majority of libraries had closed stacks, of a particular item in a library collection quickly. The way in which the call number is identified and displayed is critical to the success of a user-cordial online catalog. On the traditional catalog card, the call number is printed in the upper left-hand corner, just as it will appear on the spine of the item, i.e., in multiple lines. This is done to facilitate transcription and recognition by the user. Today, most online catalogs display the call number on one line; often the call number is not identified at all or it is cryptically identified, e.g., "CAL." At least one online catalog, the Minnesota State University System, with its online catalog developed at Mankato State, identifies the call number as a "location number," although it too is displayed on one line. Furthermore, the call number is typically not set off on a screen by special characters or blank space. In this respect, the traditional card format is superior to online catalog displays. Smug system designers should take note and get busy on devising a solution. Perhaps the

call number on a CRT screen should be separately located at the lower left-hand or right-hand corner of the screen and displayed in a multi-line format. Placing the call number at the top of the screen would mean having the patron turn his head and eyes a greater distance, thus causing discomfort and inconvenience.

Screens that provide a display of information in columns or tabular form must provide labels for each information element--including line numbers. Assumptions must not be made about the user's ability to understand unlabeled information, e.g., postings.

Another problem with most textual displays is that they use text lines of too great a width. Ideally there should be about 50-55 characters per line, assuming a possible 80-character display width capacity.[14] Further, unjustified text lines are just as legible as right margin justified text.[15] Lines should be broken at words rather than splitting a word in half.

Videotex, which uses a character generator and RF (Radio Frequency) communication techniques, can only display a maximum of 40 characters per line. This seriously reduces the amount of information that can be displayed at one time--a real online catalog design limitation. However, the possibility of utilizing the graphics and color options of Videotex offers exciting possibilities for the widespread dissemination of information contained in a library's catalog.[16]

The technology of the CRT terminal itself offers the designer of the online catalog additional opportunities to facilitate the recognition of information by the user. For example, individual characters, as displayed on the terminal itself, can be made unique and recognizable in a number of different ways. Words or characters can be made bold or can be underlined. Reverse video, bold and reverse video, double width characters, and double height characters may also be employed. For example, the call number (or perhaps the "location number") could be displayed on the screen employing reverse video. Field labels could be displayed with characters of regular intensity while variable information (bibliographic data, messages, etc.) could be displayed in bold (brighter intensity). Those few system designers working in the "Park Avenue" library environment can even provide for the librarian, not for the user, a color terminal, and display subject headings in red letters!

While in some libraries the public will need access to a few terminals that will display the full ALA character set, the display of non-Romanized characters will either require the use of special customized terminals or the use of bit-mapped displays.

Content and Sequence

The amount of information that should be displayed to a user is a question that has troubled librarians for a considerable period of time. In a

card catalog environment, the solution has been to provide all bibliographic information on the main entry card. Yet in a classic study conducted at the University of Michigan, Palmer explored the potential of computer catalogs and found that a successful search could be accomplished 84% of the time with only author, title, call number (reflecting location), date, publication, and subject heading information.[17] If contents notes were included, the success rate would increase to more than 90%.

Based on the results of a more recent survey, a display providing the following data elements would satisfy over 97% of reader and staff needs:[18]

- o Names (personal, corporate, and conference).
- o Title, subtitle.
- o Uniform title.
- o Subject headings.
- o Added entries.
- o Volume number, volume title.
- o Edition statement.
- o Date of publication.
- o Edition and history note.
- o References.
- o ISBN or control number.

In the online environment, data stored for use by library staff not have to be displayed to users. Various parts can be extracted from a complete master record for public display. Users can even be given the option of selecting the level of completeness desired. One of the principal strengths of the online catalog from the user's perspective is flexibility in controlling the amount of the information that is displayed. Thus, the motto for the online catalog should be the opposite of the current popular saying, "If you've got it (the full MARC record), flaunt it."

The online catalog can also be used to make the display of information more user cordial. Consider that most online catalogs now display all matching author entries plus the primary "different author" when the author being searched is an added entry. For example, if I did an author search for "Peters," and Peters appears to be the third or fourth author, in a single line "index display" Peters is typically not displayed.

If catalogs were to list the name being searched in the brief index display formats, even if that name is not the main entry, this would reduce the confusion that the user experiences when search results do not contain the term requested. Users of the online catalog may not know or care about the sequence of authors, especially for the brief record display. The sequence of authors should, however, be preserved for medium length and full record displays, as users will find this especially important for the preparation of bibliographies.

Transaction logs for several online catalogs indicate that a very small proportion of users (about 5%) requests a full MARC bibliographic display.[19] Thus, designers of online catalogs should plan for relatively brief displays to be shown automatically (as a default), and display more complete information at the user's request.

Given the significant amount of subject searching that is occurring with the majority of online catalogs, all but the briefest (index) displays should include the full title, subject headings, and call number.

The sequence of information on the CRT screen is just as important as the content itself. Typically the sequence includes, from top to bottom, an indication of the user's prior command or action, which resulted in the present screen being displayed; the primary information itself--be it bibliographic information, messages, explanatory text, etc.; and the prompts or options currently available to the user. Designers of flexible online catalogs have the opportunity to provide users with the information that they require in the precise sequence in which they are most likely to look for it, without regard to the prior practices of 3x5 librarianship. At least one online catalog, the University of California's MELVYL system, allows the user to specify the data elements to be displayed, e.g., subject headings only.

Vocabulary

The issue of the vocabulary that is displayed to the user of the online catalog is of paramount importance. Most online catalogs use labels to help identify the component data elements of the bibliographic record. However, some use two- or three-letter mnemonics or abbreviations to identify bibliographic data elements. Examples of the kind of truncated bibliographic jargon that is incorporated into existing online catalogs include: ME for main entry, CO for collation, IM for imprint, TI for title or title statement, SUB for subject, and CAL for call number. In addition, the same labels or mnemonics found in different online catalogs often imply or have different meanings. In some catalogs, for example, TI or "title" may actually include more than title statements (imprint and/or collation). "Call number" may include library and location within the library, in addition to the call number itself.

As noted earlier, the Minnesota State University system identifies the call number as a "location number," or LOCTN. From the patron's perspective it would appear, based on experience with this online system, that users understand and prefer the concept of a location number much more readily than a call number. Questions about the location number and its meaning tapered off after its introduction, and are relatively infrequent when compared to prior years' experiences at the reference desk in answering questions pertaining to call numbers.

In addition to the display of bibliographic information, a system interacts with users through prompts, diagnostic messages provided as a result

of a user's error, information messages, and status messages. Galitz[20] has suggested that to minimize ambiguity, a message should:

- o Use short, meaningful, and common words.
- o Not use abbreviations.
- o Not use contractions or short forms.
- o Use brief, simple sentences.
- o Use affirmative statements.
- o Use active voice.
- o Order words chronologically, e.g., "Enter search and press return" NOT "Press return after entering search."

Given the progress of online catalogs to date, self-explanatory screens are a realistic goal. Almost one-third of the users in the CLR study learned to use the online catalog by reading the instructions on the screen and another 20% learned to use it "by myself."[21]

Explanatory text, instructions, and prompts must be carefully chosen and, hopefully, be pre-tested. IBM, for its 24-hour automated teller machine, originally asked the user to ENTER dollar amounts for each transaction. After discovering some people searching for a door to "enter" the machine, IBM now ask users to KEY IN information.

Given the low literacy levels of the general population, individual words must be judiciously chosen to avoid ambiguity and to ensure that the word itself is understood by the vast majority of users.

Typography

Most online catalogs display bibliographic information in a combination of upper and lowercase letters. However, at least one, the Claremont Colleges' online catalog, displays all information in uppercase only. Research has demonstrated that people read upper and lowercase information more easily than uppercase print only.[22] However, people do read data element labels faster when they are all uppercase. A common failing of many screens is that labels (captions) and data tend to blend into one another. This problem should be easily solved through the use of spacing and uppercase-only labels.

Spacing

In books, blank space costs money; therefore its use is carefully controlled. Blank space is almost free when used on computer screens, and therefore can be used more frequently to emphasize and improve the readability of the information being displayed. People do have preferences for the amount of information displayed on a screen and these subjective readings decline if more or less information is presented when compared to the preferred amount.[23] A well designed page of printed material has a density loading of about 40%. CRT display screens that are subjectively preferred have a density loading of

only about 15% to 20%.[24] Thus, the spacing of information on a CRT screen--which typically is 80 characters in width, with between 16 and 20 lines of usable display space--must be carefully considered. Reynolds, in a study of COM catalogs, suggested that the chunking or separation of information through the use of space improved the readability of COM catalogs.[25] This is likely to be true for online catalogs as well.

Punctuation

Most online catalogs do not display bibliographic information in a consistent manner. A number of online catalogs use a slash, hyphen, colon, or commas to separate data elements within the bibliographic record. System designers and librarians must challenge their past assumptions in the system design process. The challenge is to think about the display of information from the user's point of view. Why should ISBD punctuation, that was developed for the printed medium, be preserved in the online catalog environment? Barbara Markuson has suggested we are actively creating a bibliographic petit point through the use of dots, slashes, and dash-dash. The continuation of traditional punctuation, codes and abbreviations, indentation, etc., would seem to be based on the desire to make librarians comfortable rather than on a desire to meet the needs of the user. This is backward thinking!

Based on the results of available research and the suggestions made by various system designers, the bibliographic record shown in Figure 8 is offered as one possible option for the display of bibliographic information on a CRT screen.

Color

As in the case with the layouts of CRT screens, there is little specific research data about using color in screen design. However, some useful advice may be found in Krebs,[26] Christ,[27] and Robertson.[28] Galitz also provides a chapter on color in his Handbook of Screen Format Design.[29]

Before we consider the issue of color in more detail, it should be remembered that, while color has a high attention-getting quality, about 10% of the male population (less than 1% for women) have some form of color vision deficiency. When color is used improperly, it can cause a performance degradation. And finally, while the costs for color CRT terminals are falling, they are still a relatively expensive item--especially for a library considering a large number of online catalog terminals.

Colors of most displays can be grouped into a variation of what is perceived as red, green, blue, and white. The color coding scheme must be relevant and known to the user. Not knowing the color code's meaning will distract the user. And, most importantly, color must be used consistently. Color associations already exist in the world at large and must not be ignored, e.g., red means stop or danger.

FIGURE 8

AUTHOR: National Research Council, Committee Intercity Highway, Freight Transportation
TITLE: Piggyback; 5 reports
CONTENTS: History and regulation of trailer-on-flatcar movement, by J. Ayre; The public interest and course of action in optimizing rail-highway transportation, by R.S. Reebie; Factors in future developments of rail piggyback, by B.N. Behling; Effect of piggyback operation on volume of highway track traffic, by A.C. Flott; Containerization in transporting agricultural perishables, by J.E. Clayton.

SUBJECT: Truck Trailers, Demountable

PUBLISHER: Washington: Highway Research Board, Division of Engineering, National Research Council
NOTE: Papers sponsored by the Committee on Intercity Highway Freight Transportation and presented at the 45th annual meeting of the Highway Research Board.
YEAR: 1967
Includes bibliographies

CALL NUMBER:
TE7
H5
no. 153

While the human eye can discriminate more than eight colors,[30] as colors are added to a screen, confusion and the time needed to differentiate between the added colors will increase. In general, alphanumeric displays should be restricted to no more than four colors; graphic displays may safely use more colors.[31] And the warm colors of red and yellow appear larger than the cooler colors of green and blue.[32]

Call to Action

Designers of online catalogs have the power to move the current prototype and first generation online catalogs toward a "user-driven," second generation online catalog.

To meet this objective it would seem that several steps should be taken:

1. Designers of online catalogs should stop working in a vacuum. They should meet periodically to discuss issues of mutual interest and share new concepts and techniques.

2. Designers of online catalogs should find out what users need. As noted by Shneiderman,[33] more controlled studies of online catalogs need to be conducted. Obtaining the reactions of users to alternative screen layouts and displays on a systematic basis will do much to turn the seat-of-the-pants art of screen displays into a science. Also, the results of these studies, however brief, need to be shared in the literature.

3. Designers of online catalogs should call a stop to the needless proliferation of jargon. The profession needs a standard glossary of online catalog terms, complete with definitions, so that system designers, librarians, and users of online catalogs can begin to use the same terminology when describing or identifying the same concept, screen, etc.

4. Designers of online catalogs should try to limit the number of search "surprises." Work should commence to develop a set of consistent, standard screen display formats for use in existing and future online catalogs. Much as the layout and arrangement of data elements are generally uniform from card catalog to card catalog, so should users find online catalogs; they should not be faced with the prospect of learning to decipher a new online catalog display in the search process in each different library.

5. Designers of online catalogs should not require users to learn more than one new language. Work should commence to develop a standard online catalog command language. Online catalog designers should join librarians in the standards development process.

6. Designers of online catalogs should be willing to change existing designs, however "elegant" from a programmer's standpoint, to accommodate the needs of library users.

7. Consideration should be given to the preparation of a report that presents a synthesis of available knowledge and research about the display of bibliographic information.

References

1. James Martin, Design of Man-Computer Dialogues (Englewood Cliffs, N.J.: Prentice Hall, 1973); Stephen E. Engel and Richard E. Granda, Guidelines for Man/Display Interfaces, Technical Report 00.2720 (Poughkeepsie, N.Y.: IBM, 1975); B.R. Gaines and P.V. Facey, "Some Experience in Interactive System Development and Application," Proceedings of the IEEE 63 (June 1975): 894-911; T.F.M. Stewart, "Displays and the Software Interface," Applied Ergonomics 7 (September 1976): 137-46; Stuart Sutherland, PRESTEL and the User: A Survey of Psychological and Ergonomic Research (London: Central Office of Information, 1980); Ben Shneiderman, Software Psychology: Human Factors in Computer and Information Science (Cambridge, Mass.: Winthrop, 1980); Wilbert O. Galitz, Handbook of Screen Format Design (Wellesley, Mass.: W.E.D. Information Sciences, 1981).

2. Joseph R. Matthews, Public Access to Online Catalogs: A Planning Guide for Managers (Weston, Conn.: Online, 1982).

3. F. Wilfrid Lancaster, The Measurement and Evaluation of Library Services (Washington, D.C.: Information Resources Press, 1977); Ruth Hafter, "The Performance of Card Catalogs: A Review of Research," Library Research 1 (1979): 199-222; Karen Markey, Analytical Review of Catalog Use Studies, Research Report Number OCLC/OPR/RR-80/2 (Columbus, Ohio: OCLC, 1980).

4. James R. Dwyer, "Public Response to an Academic Library Microcatalog," Journal of Academic Librarianship 5 (July 1979): 132-41.

5. H. H. Neville, "Computers and the Language of Bibliographic Descriptions," Information Processing & Management 17 (1981): 137-48.

6. Matthews, Public Access to Online Catalogs.

7. Linda Reynolds, Visual Presentation of Information in COM Library Catalogues: A Survey, Volume 1: Text and Volume 2: Appendices. British Library R & D Report No. 5472 (London: The British Library, 1979); Linda Reynolds, The Presentation of Bibliographic Information on Prestel, British Library R & D Report No. 5536 (London: Graphic Information Research Unit, Royal College of Art, 1980).

8. Joseph R. Matthews, Gary S. Lawrence, and Douglas K. Ferguson, Using Online Catalogs: A Nationwide Survey (New York: Neal-Schuman, 1983), 123.

9. Norman D. Stevens, "The Catalogs of the Future: A Speculative Essay," Journal of Library Automation 13 (June 1980): 88-95.

10. Charles R. Hildreth, Online Public Access Catalogs: The User Interface (Dublin, Ohio: OCLC, 1982).

11. Ibid.

12. Thomas S. Tullis, "An Evaluation of Alphanumeric, Graphic, and Color Information Displays," Human Factors 23 (October 1981): 541-50.

13. Benjamin Scott Fryser, The Effects of Spatial Arrangement, Upper-Lower Case Combinations, and Reverse Video on Patron Response to CRT Displayed Catalog Records (Provo, Utah: Brigham Young University, School of Library and Information Sciences, 1981).

14. Rolf F. Rehe, Typography: How to Make it Most Legible (Carmel, Ind.: Design Research International, 1974); Wilbert O. Galitz, Human Factors In Office Automation (Atlanta, Ga.: Life Office Management Association, 1980).

15. Rehe, Typography.

16. Michelle Toombs and Bob Wilson, "The Calgary Libraries Telidon Trial," Information Technology and Libraries 1 (December 1982): 331-41.

17. Richard P. Palmer, Computerizing the Card Catalog in the University Library (Littleton, Colo.: Libraries Unlimited, 1972).

18. Alan Seal, Philip Bryant and Carolyn Hall, Full and Short Entry Catalogues, BLRD Report 5669 (Bath, England: Bath University Library, Centre for Catalogue Research, 1982).

19. Ray R. Larson, Users Look at Online Catalogs, Part 2: Interaction with Online Catalogs (Berkeley, Calif.: Division of Library Automation and Library Studies and Research Division, University of California System-wide Administration, 1983).

20. Galitz, Human Factors.

21. Matthews, Lawrence, and Ferguson, Using Online Catalogs, 105.

22. M. A. Tinker, Legibility of Print (Ames, Iowa: Iowa State University Press, 1963); E. C. Poulton, "Rate of Comprehension of an Existing Teleprinter Output and of Possible Alternatives," Journal of Applied Psychology 52 (1968): 16-21; Galitz, Human Factors.

23. P. C. Vitz, "Performance for Different Amount of Visual Complexity," Behavioral Science, 2 (1966): 105-14.

24. M. M. Damchak, "CRT Displays in Power Plants," Instrumentation Technology 23 (1976): 29-36.

25. Reynolds, Presentation of Bibliographic Information.

26. M. J. Krebs, "Design Principles for the Use of Color in Displays," in 1978 SID International Symposium Digest of Technical Papers (Los Angeles, Calif.: Society for Information Display, 1978), 28-29.

27. R. E. Christ, "Review and Analysis of Color Coding Research for Visual Display," Human Factors 17 (1975): 542-70.

28. P. J. Robertson, A Guide to Using Color on Alphanumeric Displays, IBM Technical Report TR .12.183 (Poughkeepsie, N.Y.: IBM, 1979).

29. Galitz, Handbook.

30. Krebs, Design Principles.

31. Galitz, Handbook.

32. W. H. Tedford, S. L. Berquist, and W. E. Flynn, "The Size-Color Illusion," Journal of General Psychology 97 (July 1979): 145-49.

33. Schneiderman, Software Psychology.

QUESTIONS AND DISCUSSION

Comment: I'm not sure that I would buy the concept of standard language and standard format. It seems to me that there's an assumption that all of the requirements of all library users are the same. I don't think that's the case. It is interesting, however, that we do have such a thing as standardized toilet paper holders. You can go buy 35 different brands and bring them home, and they fit. But that's also because the use of that is pretty standard. (laughter)

I hear what you're saying. It isn't that I disagree that strongly, but I think that in so many cases, we are all insisting that we are different. And in some cases, we really are, maybe not greatly, but we are different.

Matthews: I don't think there's an inconsistency between what I am advocating and your statements. What I'm suggesting is that for the elements that are common among all the various systems that are represented here and not here, that we should call them the same thing. We should call it an index display, rather than the list of seven or eight terms that we currently use for index display. And define what an index display is. It's a single line for each bibliographic record. It's got a line number. It's got x number of characters in a title or call number or whatever else that you're talking about.

Just because we have standard terminology doesn't mean that each system has to employ that particular concept or screen or whatever. But it's this terrible, "If I didn't invent it, and I don't name it, then it can't possibly be good," that I'm concerned about.

Comment: I guess my disagreement has to do with degree rather than general philosophy. I don't think you can standardize something you don't understand. I think we're a long way from understanding the functions that are required in an automated catalog, much less what the command language ought to be.

We're much more likely to be able to standardize, for example, on terminology for the labels on a display screen, before we can decide what goes on particular display screens. I'm just concerned that premature standards activities are going to produce standards that nobody will see fit to follow. So that's a step backward.

Also, I wish to disagree on a small point, when you said that lines are free on a CRT screen. They aren't.

Matthews: They're relatively free.

Comment: No, I don't think so. The tradeoff is between making the screen relatively more readable or going to two screens. That's a difficult tradeoff.

Comment: You're also assuming it's a CRT. They're going to be using hard copy also. Those blanks are not free.

Matthews: But then it's a design issue. Should you design the system for the relatively small number of terminals that are going to be hard copy, or should you design it for the vast majority of terminals that are going to be CRT terminals, at least for the next few years?

Comment: In defense of what has just been said about standardizing the command language, I prefer the analogy of the automobile industry, where we have VWs and trucks and Mercedes and so forth. But each one of these has a steering wheel, a brake pedal, and a gas pedal--the user interface.

Comment: Where's the gearshift?

Matthews: Because it's either on the floor or on the column or on the...(laughter)

Comment: We are now employing a lot of synonyms for the arguments to commands. So in our system you can now say display review, brief....You can say anything you want and get it, assuming they're all synonymous, which is a point I think you were making implicitly.

That phrase someone used earlier, "Park Avenue terminals," I like the term. I have a lot of trouble with Park Avenue terminals. As a designer, unless you have the luxury of knowing what type of terminal you've got out there, the whole business about terminals can really paint you into a corner quickly. Sure, it would be nice to have reverse video on the call number. But how do you know the terminal can support that feature?

ISBD. I once wrote programs, many years ago, that could more or less bust up ISBD records. But I have enough experience to know that you could get into trouble if you tried to systematically and uniformly remove all ISBD punctuation before you displayed the record. I'm not sure we're ever going to be happy about sanitizing ISBD records for display.

Finally, blanks definitely are expensive, if you're using a relatively lowbrow telecommunication protocol. Which blank do you send out as a character?

Matthews: Let me make it very clear, I am not advocating color terminals. I do know a number of librarians, especially technical services librarians and some reference librarians, that would love color terminals.

Comment: Those are the same ones that wanted red subject headings.

Matthews: That's right. I once heard a librarian tell Fred Kilgour that the reason he wasn't successful earlier was that he was unable to generate red subject headings on his computer-produced cards.

Comment: In all honesty, as a vendor, we have been asked for color coordination on the frames of our terminals. (laughter)

Comment: I don't think we should dismiss terminal features so easily. It does make sense to take a look at each available feature very carefully, like the split screen idea, where you could be displaying terminology over here, and the LISA-type concept of windows...

I think we shouldn't box ourselves into the cheapest common denominator because those features can have their proper role in appropriate places. It would be very hard without knowing what all the uses and users are at this point to really say, "Don't do this, don't do that." So I would be careful.

Color, as another effect, can be very nice and effective. It just happens to be still somewhat expensive. I would love to use color if it's the commonplace thing, which it will be one of these days.

Comment: Just a point on the perversity of standards and the way that old catalog card hangs in there. I guess I'm not giving away any secrets. I saw a prototype system not long ago that I can't talk about; it's not one of ours. One of the features is that when they display a list of subject headings that's been retrieved from a stored file, the graphics make it look like these are each typed on the top line of a shingled deck of cards. They say they do this because that's the way they perceive that the user thinks of it. And, by God, it was attractive. It caught your eye. It was surprising. If they told me they were doing that, I never would have believed it. But seeing it...

Matthews: How about the hole on the bottom of the card? (laughter)

The reason that I'm particularly attracted to the issue of standards is that, one, if we don't start fairly soon, given the time it takes to develop standards, nothing is ever going to get done within a reasonable period of time.

The second thing is that there's only a few online catalogs now. Five years from now, there's going to be a heck of a lot more in the way of online catalogs. And if we, at that point of time, somehow come up with the beginnings of a standard, there's going to be a tremendous amount of resistance on the part of IBM and Geac and CLSI and DataPhase and Data Research and everybody else out there who has spent a lot of money developing code and has a greater installed base. There's going to be much more resistance at that point in time. If we start now while things are relatively small in terms of what the financial implications are going to be...I'm also

not expecting that the initial set of standards is going to be a one time only kind of process. I view the standards process as an evolutionary kind of process.

Comment: I do believe that we should get started right now. We must be very careful, though, because there are so many differences, just to reinforce what was said earlier, that we must be certain that we are not applying the standard terminology or standard command to things that are, in fact, different. There are enough similarities, but yet enough differences, to cause us some problems there.

Comment: I think standards are two-edged swords. You know, swords for us to fall on. They do provide a fixed target for everybody to aim towards. By the same token, they represent an inertia in the evolution of ideas. If we develop a standard, everybody seeks to achieve that standard. We tend not to branch off into other innovative technologies and approaches.

I think the card catalog itself is a real good example of that, an incredibly stable standard that nobody, because it is a standard, has thought to evolve. In fact, the standardization of the physical form of the catalog card is still trying to find its way into the automated world. And I think it impedes some of the more innovative and evolutionary thinking that we should be doing. So I think we need to be real careful about how we look at standards. There are benefits, but it may come around and bite us.

Matthews: And one of the things clearly in terms of the standards process...labels for information on the screen. The data elements themselves ought to be called what users call them, and not what librarians call them. Then you get away from "entries" and "imprints."

The best story I ever heard about a name entry was in an article that appeared in Reader's Digest. The patron went to the card catalog and looked up a reference, and it said, "see main entry." So he left and went out to the main door of the library and he's out there, looking around... (laughter) The librarian came up to him and said, "What are you doing, Charlie?" He said, "The card catalog told me to see the main entry. Is there something special i'm supposed to look for?"

So it's the jargon of librarians and system designers that we've got to make sure that we sanitize when we present that information. But a part of the standards process is also saying, we don't know whether authors should come first or titles should come first, or whatever should come first. And what's the appropriate sequence of information? That whole issue is not very well understood yet.

Comment: It strikes me that one of the problems that we have in trying to solve some of those issues has to do with the difficulty of conducting really controlled studies, because systems are not portable. If you sequentially try out two systems on the same population, they're going to

be conditioned by what they saw the first time, which will make the results of the second test less valid.

You see the problem. I don't know what to do about that.

Comment: To say nothing of library directors who refuse to expose their patron population as experimental objects. Not so much that they already have a system in operation, but to change that system dramatically for experimental purposes is not very acceptable.

Comment: What you'd like to be able to do is erase the memory of the population every time you try something. (laughter)

Comment: In an experimental situation, as a researcher and wanting to do more, there are some practical problems. But they're solvable; there are standard techniques. You're talking about controlled experiments. You don't really need a huge sample, 30, 40, and 50. I won't get into sample selection.

We're doing it now with different interfaces. I have dial access too, so there are limitations of baud rate and so forth. My problem in being able to run subjects in controlled experiments to get answers to some of these kinds of questions is really very simple. And there are standard techniques to deal with the fact that they're biased by the ones that you switched among. I don't want to get into all the research technique jargon. Those problems are solvable.

What I need is the sanction to do it and the money to do it. But, really, the problems you're talking about aren't that great in running subjects in a controlled environment, finding answers to whether they want labels or not, whether they want uppercase or lowercase...

Comment: You're in a unique situation. You have dial access to five different systems. And a given population at institution x doesn't.

Comment: But how did I get it? Through the generosity of the people around the states. I've got access to around 30 systems. I didn't get that permission from God.

There are problems that are solvable, but the real problem is getting people to support that kind of experimentation. It doesn't have to be completely centralized.

Comment: Yeah, but it can't be conducted on a wide scale by most of the people in this room.

Comment: Isn't that true, though, that what you're doing is dialing into 30 systems. It's not that your sample size is so small, it's that each of those 30 systems represents a tremendous investment by the people who have

delivered and created that system. So if your results are that systems 7 and 12 are the best, that doesn't necessarily convince the other 8.

Comment: I'm not evaluating systems.

Comment: I didn't say you were. But my point is that that doesn't persuade those 30 people, those 30 system directors, that they have to go out and bring the capability in here.

Comment: Responding to that point really leads me back to the question that I've wanted to raise for some time. And it's why I feel a little bit less enthusiastic than you, it's a matter of degree, about the formal standardization process.

It was said that it's a two-edged sword, that standardization can either drive uniformity or drive rebellion in uniformity, and it can also thwart creativity and so forth, the same point you made. I think you could get a consensus around this table that the result or the end, if you like, of standardization, which we have t keep in mind, is only a means; you mentioned what it was--two things, greater uniformity and consistency. So let's take those two things, uniformity and consistency.

Standardization is one way of driving towards that, but there are these other problems. It's a triple-edged sword, not just a double-edged sword. I think we then seriously need to answer the question, because there are a lot of reservations among us about it, what other means besides formal standardization efforts might get us, in short, promote, get us closer, evermore, slowly, gradually, to that uniformity and consistency.

Conferences like this are one way, just sharing and knowing how they did it somewhere else. I'm not sure formal standardization procedures and all the sanctions behind it are the only way to achieve the end. There may be equally desirable, perhaps better, ways of going about it.

I'm asking this question. You said nothing is going to get done, and so don't do this. I think that's putting it too strongly. What alternatives are there to move towards that goal?

Question: If you look beyond standardization, would you have any information on the usefulness of specific techniques like use of touch screens, or psychological factors that the man-machine literature points out, such as if you use a menu, don't put more than seven choices on the screen because memory just can't handle it? Practical pointers that this group could benefit from.

Matthews: That was one of the calls for action, to put together a synthesis of those kinds of things. Certainly, based on the CLR data, the patrons in the CLSI libraries equally loved the online catalogs with the touch

screen as opposed to keyboard. The one frustration for people that only have the touch screen is that as you become proficient, it takes a while to step through that. If you make a mistake, you've got to go back x number of screens to get started down a different track.

I'm an advocate of multiple technologies, because there are multiple needs out there. Touch screen only is not the answer; keyboards only is not the answer. I think a mix of technologies is going to be real important.

Comment: The other thing that emerges for me as I hear you talking about standards and hear reactions....It isn't so much that the end is uniformity. It would seem to me that as we get more information back, as we find out how people can use an online catalog more easily, that helps us in design, which may somewhere down the line make them more uniform. If nothing else, it ought to make them easier to use so if you go to somewhere where it's different, you still will be able to use the thing.

I think as I look at it, and as we struggle with our design, part of the problem is that I don't have enough information. I don't have that feedback. So I'm not as concerned with standards as I am with those things that I know will help, which in and of themselves, if we listen, may lead to a kind of uniformity.

Comment: The changes that have been suggested are things that affect two programs and a couple of code tables. That's less than a couple of man-weeks' work. It's easy and trivial. I only regret I didn't know that a long time ago.

Matthews: In some respects, I think it is easy if the design at the outset is set up to establish flexibility. Unfortunately, not all of the designs of the online catalog are table-driven such that you can change the format, move the information over, down, take it off, put another on, change the label, etc.

Comment: I think the difficulty is that all of us are so often involved with deadlines and operational problems that we haven't built this into our own way of working. It's a point that Charles was getting at this morning, the whole design process, lax commitment to an experimental element.

Shneiderman states the case so strongly that if I were following his precepts I would test everything with a controlled experiment before I could do anything, in which case we'd have to pass a law delaying the 98th Congress for about ten years. We've already requested that they not elect members with the same last name. (laughter)

Nevertheless, I think that we can ourselves, in our own design procedures, grab a hold of some element in the experimental processes and use it when we can. It doesn't always have to be a simulation of the system, by the way. Paper and pencil kinds of things and screen layout designs, for

example. There's a classic example where you don't have to put one of them on line to have some feeling for which is more readable and understandable.

One more point I'd like to make, and that is that in this issue of standards and the experimental process, if you gave me a choice between spending some money on some research or spending some money to support some group working on standards, I'd rather have the results of some experiments than I would the collective ignorance of all of us working on standards.

Matthews: I think the point about alternative ways of arriving at the same goal is a good one. Maybe the standard setting process is not a--what I'm saying is, we have a goal that we're trying to strive towards, and I'm trying to get us started down the road.

Comment: Another way to look at the standards formulation process is to look at it as wringing out all the frivolous deviation. That's sort of the first step. It's a slightly different way of looking at it, as opposed to saying uniformity.

Secondly, there are two ways to arrive at a standard. One is the evolution towards consensus. The other is serendipity. Serendipity is when people look at what other people are doing, and the things that don't work slowly but surely get wrung out.

That's sort of two ends of the spectrum. Wherever we wind up on that spectrum, it seems to me the communication is critical to getting there, and these kinds of meetings facilitate getting there. I hope this meeting is not going to be the last one.

Question: What about guidelines in this area as opposed to standards? That is, broad guidelines, taking the first step. I think that's what the discussion is pointing at.

Matthews: There is a body of knowledge out there now in terms of system design, how to display information, that could be synthesized and put together for designers of online catalogs and bibliographic displays, which I think would be quite helpful. That could be taken a step further, once that document had been prepared, in terms of passing it around to you for your reactions, making modifications, etc. I think that would be quite a valuable document, in terms of some general principles. I think we've been talking about one of the real underlying principles, which is to make your system easy to change.

The second step of that process is to take the Benjamin Fryser study the next step. He looked at two formats. Let's look at 20 or 30 formats in comparison to one another and experimental situations, and pick it up on a microcomputer, cart it around the country and expose it to a lot of people,

and gather some data relatively cheaply. That would provide an awful lot of information.

Question: Is there any evidence that would suggest that the users themselves, if they were led through the process, would be willing or able to create their own display format as a profile, if you will, for them as individuals, so that the next time they want to come up and invoke that profile or change it, that they could do it? Is there anybody doing that?

Matthews: Not that I'm aware of. I think it would have some incredible machine implications.

Question: But would users use it?

Comment: The eighth-year graduate students. (laughter)

Comment: Even with editing systems, online editing systems used for programmers, they take the defaults, whatever the computer center set. Very seldom do they do anything to fix up that profile.

Comment: We have some experimentation with user profiles. There are a couple of problems. First of all, to benefit from the profile, the user has to be very proficient with the system.

Question: These profiles, are they display profiles or other types of profiles?

Comment: They're all kinds of profiles, including display, so that when the user logs on, the system is configured just the way he likes it, with all his defaults, rather than all our defaults.

It's a serious problem to do this on a wide scale; you get into a user identification problem. It's antithetical to being user friendly. You have to identify yourself when you're coming on. But for a few reference librarians we've actually done it, because they mandated it.

Question: Maybe I didn't understand you. Were you saying that the studies that led up to the recommendations on the last page were not that solid or definitive? We should not accept this as the best recommendation?

Matthews: A lot of the information that is out there is basically kind of intuitive--based on my experience, here's a set of rules and guidelines. For example, Wilbert Galitz wrote a handbook on screen format displays. It's great; it's really super. But it's not based on experimentation. It's just, here's how I done it good in my life insurance company and now I'm a consultant and...

The other thing is that a lot of the studies that have been done are one-time studies that have never been replicated, even on a small scale, to verify them and make sure that advice is appropriate for a diversity of audiences. It may be appropriate for Great Britain, and may not be for Australia or Canada or the U.S. or the South or the North or us Okefenokees out in California.

Question: If one is in the position of needing to rewrite the program that does this display anyway, would your conclusion be that labelled fields would be a better way to go, or should one hold off and make it look like a catalog card for the time being?

Matthews: Don't make it look like a catalog card.

Comment: That's the way it works now.

Matthews: Don't make it look like a catalog card. Based on all the evidence we have so far, that's the best advice I can give you. Labels should be uppercase. They should be separated with some space before the text--which is variable, which is upper and lowercase. The width of the lines should be about 55, maybe 60 characters, given the free space on the CRT.

Comment: Most of the studies of catalog use show that nobody ever looks at the whole record, 10% or less. They look at the top three lines of the card, and they're off to the stacks. I see less commonality among systems on the short display. That's another concern to me, to find out what data should be in there and how it should be formatted. I don't think you really addressed that topic.

Matthews: There are two studies. There's the University of Michigan study that identified some specific data elements that would satisfy about 90% of the transactions.

The other one is this more recent British study, which has a little bit longer list, which satisfies 97%.

Comment: Both of those I would call sort of medium entries rather than short. My question really is about short, something you can get on one or two lines. A lot of systems have one- to two-line displays.

Matthews: There is no research that I know of about one- or two-line displays.

Comment: We have four formats. We have a very short one, so you can get a lot on one screen. We have one that looks like what you find in a card catalog, but not under the main entry. Then we have a lengthy one tagged with our names, our labels. And then a MARC format.

You also can make your own with our system. You can say display title and subject, which is one that I love to use when I'm just rummaging around. If I'm trying to find some rare books by Arthur Conan Doyle, I can say, display editions and just give me the editions statements, so you can zero in on the first and second editions. But the use of those is down in the grass, as they say in radar.

Matthews: And most of them come from the designer's terminal. (laughter)

Comment: All of the talk up to this point about display formats is about bibliographic displays. That's all fine and dandy, but what we see in the requests coming in most often are minimal bibliographic information combined with holdings, location information, acquisitions information, search status. A view of the data regardless of what files it's coming from or what format it's in, that is getting to be much more in demand. Let's go to one screen. Let's get a lot more there and not just pick on bibliographic form. That gets real tough. You start taking holdings, especially detailed holdings, how do you summarize them and capsulize them onto a screen and make it readable? That's where it falls all apart.

Question: Have there been any studies with the four levels, where you take a group of people and change the default level? Is it that people don't want additional information, or they don't bother to go beyond whatever the first default level is?

Comment: We looked at that through our transaction logs. It's very rare for the user to bother changing the defaults, which says something about our study results. Did we get the results we got because of defaults we set?

Comment: We've seen two kinds of trends. One is that almost always if there's a default, that's what they accept. A related one is that if you prompt them, a gentle suggestive prompt, they almost always follow it.

Another thing I worry about with standardization is the display format. There's an area of innovation for a tri-level screen where we put related things, status information, guidance prompts, structure, whatever, error messages.

When they're looking at a single bibliographic record, right after the entry you may have three or four subject headings. Right there in parentheses after the labelled subject, somewhere embedded in the record, put: "Search on these, they may retrieve related items." That's going to be effective as hell, because that's a prompt, and they do use prompts. And that's where embedded guidance is useful.

Question: Should there be numbers on any of these fields they can search on, so that you minimize input if they wanted to chase the tracing?

Comment: We're going to have to pay a lot of attention to standards. We've been talking a lot about it here. I've heard some talk against it, but it's a positive thing. We'll work towards it. We need to recognize that many of us sitting here also are representing a very powerful force that works against standardization, and that's competition.

Comment: I wanted to follow up on the concept of whether a standard is a guideline. It's been brought to mind by this comment about competition. Probably, when the work started on what's now known as the open systems interconnect standard, no one thought that anyone would be willing to standardize that. And what they standardized is what they now call a reference model.

There's an issue that comes up in the standards proceedings of, can you standardize something that's not deliverable? And, in fact, the U.S. nearly voted no, and at various stages did vote no, on the ISO open interconnect model because they said it's not a deliverable, it's a reference model.

It seems like the standards community is now moving to the point of allowing the word "standard" to be applied to a family, an approach to developing standards. I think the open systems interconnect is a good example of a highly competitive area where a standard was brought forward at a very early date. And it has helped. There is not one standard at each level, but people now say this is a presentation level protocol. Well, that's a bad one. (laughter) The less controversial levels, down around transport and so forth, there's more than one. And it just clarifies what it is that a person is proposing. Perhaps that is in the spirit of the standard as a guideline.

I think user and consumer groups can bring over those standards without closing off the innovative options of the producer groups. Whereas if consumer and user groups bring forward specific standards of command language nature, etc., then producers do find themselves painted into a corner.

So I think that the OSI model is a good model for early standardization. However, if you know something about the politics, maybe it's not a good case to bring forward. But it's, at least, not an example of where standards came out years after the fact. They came out months after the fact. (laughter)

Matthews: Maybe we shouldn't call it a standard. Maybe we should talk about display objectives or goals or guidelines or something to that effect.

Comment: True, but in terms of literary warrant, the fight has been fought. The ISO has allowed the word standard to be applied to that document.

So in terms of literary warrant, the word standard now covers reference models.

Comment: With regard to standards, I think it behooves those of us who put them together to take some responsibility. And in the long run, it will ultimately be in all of our best interests to adhere to them.

I want to just comment on one particular standard that is particularly irksome to me. In regard to the serial standard, summary serial holdings format, if you expect the patron to read that, all I can say is, good luck. If you think that this is unreadable, just try presenting users with one of those. And near as I can tell, that's what was intended. It's highly irresponsible.

Comment: It's been said a lot of ways, but I think perhaps it would be worth stating it very concisely, that the goal is comprehension, not uniformity. If things are not uniform but they're all comprehensible, we've achieved what we want to do. That means expunging librarians' jargon. "Tracing" goes back to the clerk who traces a card out of a card catalog. "Call number" goes back to closed stacks; we haven't called for books in a long, long time in most locations.

Comment: In addition to the element of competitiveness mentioned, one of the ways in which library jargon gets into our displays is when we involve the library staff in the design process. We did this because we felt it was very, very important. One outcome of that is that the field that I recommended be called "publisher" and provided some sample screens for and so on, the committee couldn't agree on that. So it's called imprint, which I think makes real sense to the entire committee. (laughter) With maybe one or two dissenting voices. (laughter) I think it will cause no significant problems to our user community because they will look at it and see the publisher, place, and date quite clearly there. And we may be educating people to use the proper language to describe a book.

Comment: The words God gave us. (laughter)

Comment: The words we developed over hard years. But, anyway, there's that factor. There's a sort of group dynamic involving the staff in the design process.

There's also the desire for institutional uniqueness. We encouraged our staff to look at MELVYL very carefully, and to rip off as much of MELVYL and several other systems as we could. But some comments that were made in this process indicated that some people felt we didn't want to be just a MELVYL clone, that Socrates needed to have its own unique character. And once again, I'm sure that people who felt that were absolutely right.

Comment: One of the things related to screen design, and you get into real conflicts here, is that everybody knows, you display things in tabular

form and it's easy to read. So help me God, one label in our system says "physical features," and it doesn't line up. We couldn't agree with our customers on a variant form of label that took less than 20 characters, roughly what "physical features" takes. And we weren't prepared to throw away that much of a screen dropping off the title and subject headings. So physical features doesn't indent. Now, we'll happily let you come up with an abbreviation that's shorter. But we couldn't get people to come up with a variant that was less.

We're involving not the users, we're involving the customers, the library staff who are making the decision on the system. Remember, they're still not the users. I don't think we would have had that problem if we'd involved the users. In fact, I have a suspicion that they wouldn't even have put the field on the screen.

Question: Who cares whether it's 1970, right?

Comment: Only if it's filed differently.

Question: And who cares whether it's called publisher or imprint? The people in this room care, but I just raise the question, do any users care at all?

Comment: Okay, except in a few other cases where you don't know what to use, such as added entry. I wish I knew what to call an added entry so that it would be meaningful to people. I don't have any ideas. It's difficult at times to come up with the right thing. This is what I'm looking forward to, in terms of some kind of research. I think the feedback that we get may help in as simple a thing as naming what an added entry is.

VI. COMMAND LANGUAGES AND CODES

Michael Monahan

When we think of command languages for an online catalog, we face two conflicting views of its users and their priorities. One view is that users are unfamiliar with the computer catalog and need every encouragement to use it. When I think of this type of user, I tend to think of my grandfather: a nice person not even remotely mechanically inclined.

The other type of user is sophisticated and eager to make the system jump through hoops as he or she plucks exactly the right record from the complex choices available. Of course, the models for this type of user are us: the systems designers.

The obvious problem is that both types of users demand different solutions. Therefore, when we talk about command languages, we work under the stress of these two pressures. Figure 1 shows these tensions.

In this paper, we will explore the relationships presented in this figure and offer some justifications for them.

Menu-Based Solutions

We start by looking at menu-based command languages. Figure 2 shows the critical elements of a menu command structure. Since the target of a menu is an inexperienced or casual user, it is a truism that the system should be "easy to use." This means that all options must be presented directly to the user. The user commands the system by selecting from the choices presented. These choices lead the user step-by-step through the system to retrieve available information.

In order to make the system attractive, we keep the commands short: typically one character or digit. This means that one or, at most, two keystrokes are needed for each choice. More characters act to defeat the whole purpose of simplicity without offering much in the way of benefit.

In this context, we can consider a touch-sensitive terminal or a mouse-based terminal, such as LISA, as variations on the menu approach. These devices are simply attempts at better ways to select from the menu. Where the naive user is the prime target, such devices have a major role and new variations will continue to be introduced. Clearly, this is an area of ongoing research.

FIGURE 1: How does the user want to command the online catalog?

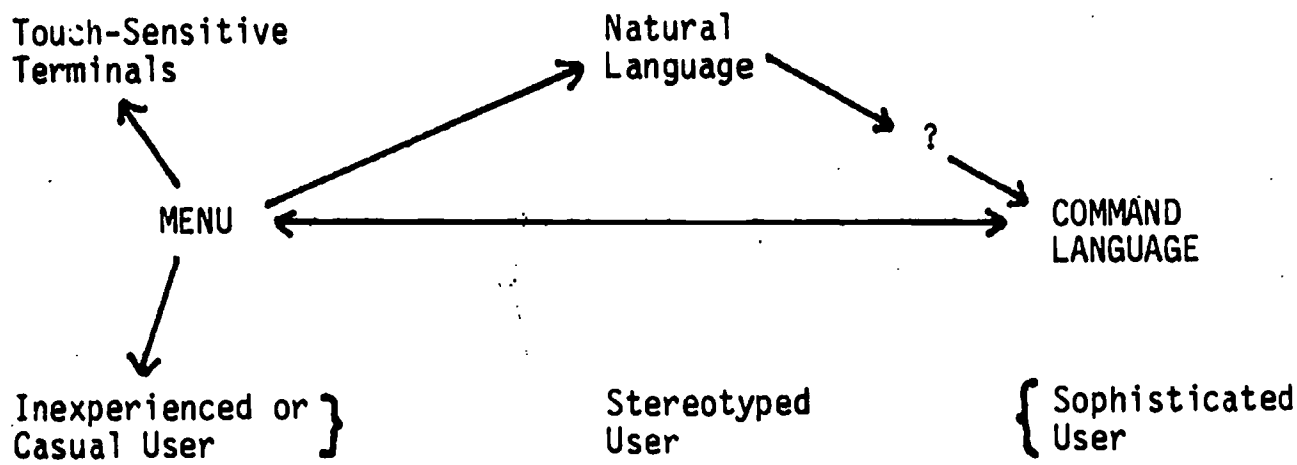


FIGURE 2: Elements of a menu system

MENUS:

--Easy to Use
--All Options Presented on Screen
--Step-By-Step
--Single Digit/Numbered Choices
--Touch-Sensitive Devices
--Natural Languages (?)

But to answer the question implied by many system designers, there seems to be little motivation to standardize such menu approaches. If a good menu system is easy to learn, what advantage is there to having one standard? Further, there appears to be little to make standard: no user should remember that choice one is title, or that six inches from the middle of the screen is author. So let us continue to experiment with menu systems and only worry about standards here as a last resort.

Natural Languages

An interesting alternative to the menu for the naive user is "natural language." This approach lets the naive user formulate a question, of whatever complexity, in his or her own words. The computer determines the intended meaning and does the necessary work. If the user has an ambiguous request, the computer detects this and prompts the user to clarify the question. The computer can also respond to requests for help, or detect when available choices should be listed.

Clearly, this is heaven for the user and system designers will be out of jobs. But, examine this sentence:

The firemen met with city council to discuss their salaries.

Clear and understandable? Not unless you know the relationship between the firemen and city council. One can invert "the firemen" and "city council" and still understand the sentence. But this puts a strain on the computer, which must understand the real world in order to correctly determine the meaning of the sentence.

So while great strides are being made in natural languages, and a significant research project in this area would be well in order, we should not expect natural language query systems for our online catalogs to be ready tomorrow.

Command Languages

This leads us to look at command languages. The first issue is: are they necessary? Consider the user approaching the catalog with this question:

Do you have a French film on computers, done since 1979, available immediately in this location?

This is not an unreasonable request, and in a large research library it is not hard to imagine queries of this complexity being brought to the system frequently. In fact, as online catalogs become union catalogs of multiple locations, media, and collections, the need to be this specific in order to limit results to manageable answers will grow.

It would also seem clear that a menu approach would be tedious and confusing for such a query. Figure 3 shows seven steps to get access to this information. Most systems would not prompt the last five criteria, and either would assume all languages, all media, etc., or would require that the operator use a special command sequence to specify these steps, if available at all.

FIGURE 3: A menu sequence for a sophisticated query.

Do you have a French film, on computers, done since 1979, available immediately, in this location?

(1)	Type of Search:	Subject
(2)	Subject of Search:	Computers
(3)	Media (if any):	Film
(4)	Language (if any):	French
(5)	Date Since (if any):	1979
(6)	Location:	Here
(7)	Availability:	Immediate

I think this example points out why we want command languages. It also shows why they will not quickly replace menu systems for the naive user.

If we accept the need to have both a menu and a command language, then what should be the relationship between the two? In many cases, the command language is independent from the menu. They are two different worlds. This would be fine if our two models of users were also independent. However, life is not so simple: the user wanting the aforementioned French film is quite likely to be our naive user. Maybe he wants the film to learn the command language!

If we keep the command language and menu separate, then any user wanting to go from one to the other must make a quantum leap. It would seem better to have the menu lead into the commands and, perhaps, teach the user the concepts embodied in the command language. If this can be done, the user learns, or the computer adapts, to the complexity of the system or user needs.

How to do this is left as an exercise. At this point, we can only make the suggestions and begin to experiment with solutions.

Standards

This leaves us with the issue of standards. At the Wye conference 10 months ago,[1] I felt that it was too early to expect standards. Coming to this

session, I was impressed that almost 50% of us raised the question of standards for online catalogs. Maybe the time to start is now.

The biggest (non-political) problem is that systems may not be conceptually compatible. Even MARC-based systems may have radically different meanings for an author search. Are series included? Are search techniques keyword, left-to-right truncation, or some algorithm? Which stopwords are used?

This incompatibility we cannot change, so let us accept it. Assume that local features will exist, and that not all systems will support all functions. We can still define a reasonable low, but not lowest, common denominator. A system could meet the standard, but would be free to extend it. While it is true that this might appear to restrict experimentation, this is a risk I suggest we take. Further, we can assume that some systems will implement the standard as a "side" module, independent from the rest of the host system. This at least offers the naive user from an external site a way to use the systems.

It might also be best to accept an imprecise standard, such as: "This is the way to do an author search, but what an author search means will differ from location to location." An imprecise standard would permit us to both agree on something with some dispatch and offer a good chance that systems will quickly meet it.

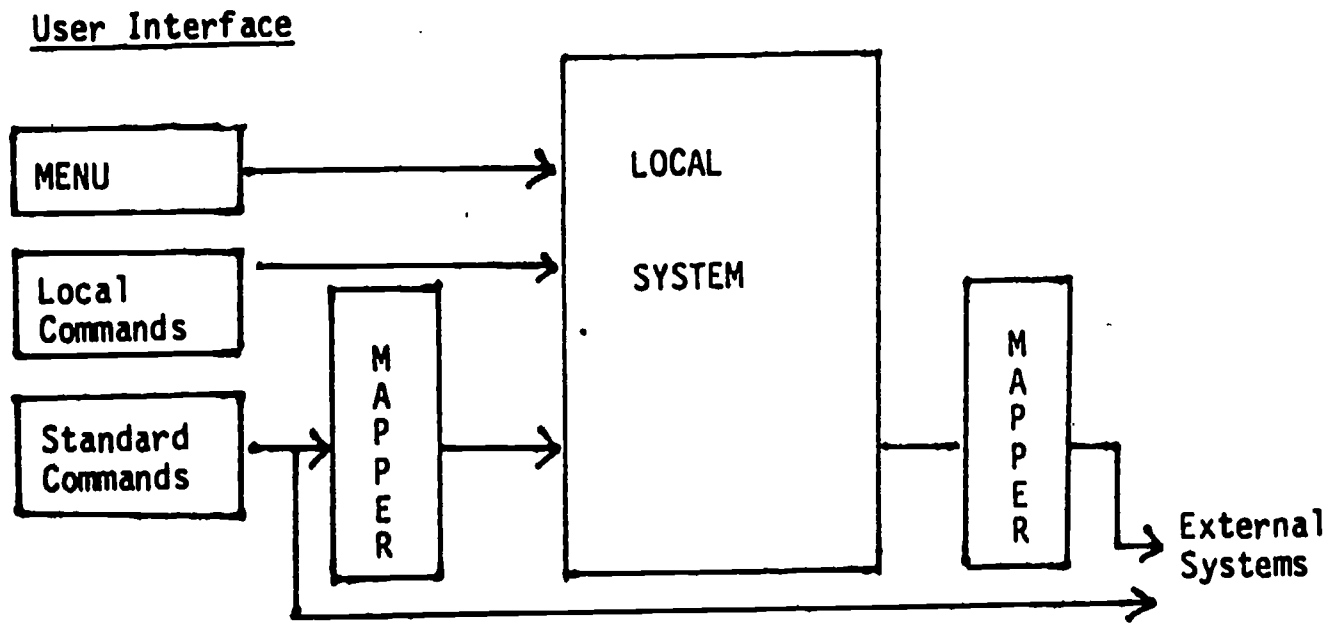
Merely because a standard is imprecise does not prevent it from being useful. When the car in front of us flashes red or orange, square or rectangular, single or double or a series of lights, we still know it is about to turn.

Figure 4 gives an outline of a system with such an online catalog. It implies that there may be up to three patron interfaces: menu, local command, and standard command. Two of these, menu and local command, go directly to the local system, while the standard command interface must be mapped into the local functionality. On the other hand, the standard command language needs no mapping to the external world.

A major point is that figure 4 is the most complex view possible. A system developer need only implement the standard command language to both have an online catalog and talk to the outside world. Local menus and command extensions become options.

I accept that there is a conflict between calling for a menu to command growth path, no menu standard, and a standard command language. But the fact is that we need to address what are different user audiences. And I suspect that the library world will quickly see the benefits as users react with genuine pleasure to being able to transfer their skills from one system to another.

FIGURE 4: Online Catalog to support standards



Turnkey library vendors do not get purely pleasant social letters from their clients as often as they would like. Therefore it is worth mentioning a case several years ago when a customer forwarded a compliment received from a user. The user had just returned from a university in the United Kingdom where he had been very pleasantly surprised to discover that they used exactly the same version of Geac's online catalog. He wrote to express his pleasure and implied that this similarity saved him time in using the system. When a library permits a scholar to preserve skills while pursuing research across multiple libraries, then it is doing its job.

I would contrast this with the situation with another familiar tool. As I travel, even among Geac offices, I am frustrated at the different variations that we seem to have found in the telephone. It's not enough that they ring, chirp, and warble, but one needs an instruction book to use them.

The library world can avoid this frustration if we have a core standard. We must accept local initiatives in delivery devices, menus, and advanced features, but we can provide a common core that will work in all locations. This, I think, is what our users will most demand from us in the next five or ten years.

So much for the technical issues, what about the politics? I am struck that if even half of the systems present here adopted a common format for common features, we would have our standard. If we simply acted to define

(1) how we now do "c mmon" searches and (2) how we would propose that they or variations could be made into a standard, then we might have it.

The biggest legitimate objection is that we do not know enough to cast things in stone. Certainly, I would react negatively to a standard that limited my ability to experiment. This is why a partial or imprecise standard is needed. This will not eliminate local change or features: if it turns out to be clumsy it can be ignored. Under this guise it is not a threat to any commercial vendor's competitive instincts. I think that there is enough to compete with that this can be ignored.

So, let's do it!

Reference

1. Davis B. McCarn, comp. and ed., Online Catalogs: Requirements, Characteristics and Costs, Report of a Conference Sponsored by the Council on Library Resources at the Aspen Institute, Wye Plantation, Queenstown, Maryland, December 14-16, 1982 (Washington, D.C.: Council on Library Resources, 1983).

QUESTIONS AND DISCUSSION

Comment: I hate menu-driven systems. I hate them with a passion. For example, SPF, which runs under TSO and CMS, where you have to wade through all those damn screens to get to the thing you really want to see. That said, DOBIS is a menu-driven system. And I still say the same thing. I hate going through and watching all those scréens.

However, we have a system that adjusts itself to the sophistication of the user. That is, when the user knows what the next question is going to be, he can answer it now. He doesn't have to wait to see the screen. He never sees that screen. And if he knows the questions that are going to be asked in the next ten screens, he can answer them now. He can answer all the questions for which he knows the answers, and then he gets to the screen that he really wants to see. It works very fast. This is a system that adjusts itself to its users. And the user knows what he's doing. No problem if he wants to go ahead and go through all the screens. Each screen gives you all the information you need to answer that screen. All the answers are on the screen.

The point I'm making is not so much that our system has solved the problem, it's just that there is a solution to that problem in that way. It can be done. I don't think that you have to go from one channel into another channel. I think there is another way to handle that.

Monahan: I deliberately tried to stay away from discussing how we might solve some of those problems. But one concern I have is that there is a difference between command chaining menus and a true command language. One of the differences is one of context. You can solve it by staying within the context of having menus, but it depends on when commands are legal and what you permit them to do. If the commands that are being chained are choice 1, choice 5, choice 15, choice 11, then you're fixed in; you've got to know exactly what the sequence of screens is and exactly how they fit together.

There are alternatives that involve chaining some sort of mnemonics together, but even that has problems. I'll let you in on a weakness, although our competitors can use this against us. Our new online catalog, in fact, lets you chain mnemonics. But it also means, since CAT happens to be one of the prompts we use, that you've got a problem looking up a subject heading called "cat" because the command decoder says: "Wait a minute, that's a command."

You've got a generalization problem there, and we've taken one choice on that. I think it points to some of the problems of locking the user in, since you're faced with how you detect that this really is what he wants to do.

Comment: The reason we need commands is because we can't use a menu system to do a search as complex as the one that you discussed in Figure 3. Suppose you start with "computers," and you say I want all items in the catalog about computers. You get a number that is a few hundred or a thousand or something, much too large to deal with. You take a look at the rest of the fields and you say: "Which is going to be the best limiter?" Say films. You say: "I want the type films. I don't want monographs, I don't want periodicals, I want films." So it goes through and says: "I have 300 films on this topic." You can limit it by language or by date or by location or by status. All of those are possible in a menu system. I can show you a menu system that can do all of those now.

Question: But how many steps is the menu approach going to take? How many responses to the machine? How much man-machine interaction was there to get that job done rather than typing in some command language?

Comment: Well, as a matter of fact, for the one she's talking about, you can put that in in what you might consider to be a command--in the way that you want it. It's just that our delimiters don't have to be blanks between the pieces of your command. They happen to be slashes. In fact, they're locally choosable characters. It looks pretty much like you want, you say this and that, or this limited by this, and so forth.

Comment: What's interesting about putting that in as one long chain is that you, as a user, don't see the development of your command. You don't see the effect of those various things. You can't use your knowledge about the system to help you. Whereas, if I go step by step, I search computers and I limit by films, and I come up with five films, I don't have to go any further. That takes care of my problem. And that probably is the case. Very few libraries have a thousand films about computers in French.

Question: How would you construct your search using command language for that?

Monahan: Well, that's a good example. In fact, one of the things that you can do to put that together is have a vocabulary of terms, subjects, languages.

Question: Then what would you type in?

Monahan: That sequence.

Question: All those characters in Figure 3?

Monahan: No, of course not, and that's the whole point. Once you move over to command language, you start to do those horrible things like get into abbreviations.

Comment: Yes, but give us the chain with the abbreviations. Make up one. SUB=computer?

Monahan: "SUB=computer/" is not a bad idea; "MEDIA=films/LANG=French (or FRE)/DATE greater than 1979."

Comment: So you would type in everything on the right and some syntax and punctuation. Maybe I'm hung up on the definition of menu. I don't like this alternative, menu versus command. I don't see it.

Monahan: What we're trying to do, I think, with command languages is to address problems of searching. I deliberately stayed away from Boolean searches, where I think the problem gets a little bit more serious.

In order to solve it, in order to present the command language to people, you're asking them to do a fair bit of typing. Nobody who is afraid to do typing is going to want to do command language. You can get through it a menu at a time. But if I deal with the real world system and they have five-second response time between steps, I can type that in in less time than it's going to take me to step through.

Comment: That's my point. There are a number of techniques between traditional menu selections, and numbered lists, and a formal command language --such as the example you give. Right there is one staring at us. And there are several other techniques in between.

Comment: The other assumption is that all those appear as separate screens. But, in fact, it is possible to have the thing virtually the way it is. And you're going to type in your search with an S, you're going to type in--and you can use or not use any one of those in order to conduct your search.

Comment: That presumes those exhaust the list of factors.

Comment: All I'm saying is, it could be 10, it could be 20, I don't care what you put on the menu. All I'm saying is that it is possible, if you have 50 different kinds of commands like that, to do searching. It is certainly possible for many of the searches that do not involve 50 commands, but just a few. There's nothing wrong with the menu approach if all you're doing is filling it in, which would be exactly the same as if you had to type something and learn a command language. That's all.

Comment: If that's all you're going to provide, a single subject, media, language, date, location, and availability, and you're going to assume an AND relationship between those, you're in fine shape with a menu. It's easier than command mode. But try to come up with a generalized menu that encompasses all the possibilities, and you're in big trouble.

I have a more general comment, though, and that's to do with my experiences of late using personal computer packages designed for end users. I think we have a lot to learn from the folks that do that. I've been uniformly impressed with the ease, particularly given my stubbornness in reading the documentation, with which I've been able to use those things with absolutely no prior understanding of how the package works.

Monahan: So what is it about the packages that leads to that?

Comment: Well, a couple of things, I think. In general, they're menu-driven. But they're menu-driven in a way that I think is only possible with a dedicated processor directly attached to the keyboard.

For example, menu choices--you can hit space and instead of having the cursor go over, each item in the menu is turned on in reverse video. If you don't like doing that, you can hit the first character of your choice. It never appears on the screen, but it's immediately picked up.

The business of keying menus to commands--I've seen this employed also, and it's essentially what you describe. You can either choose the next level of the menu hierarchy by typing the first character, or you can type three characters in a row, anticipating what you want. For example, if you want to save a file, you know that's in the extended command set. So you say XFS; that's the command for saving a file. If you just say X, however, and let about half a second go by, it will display the menu for the extended command set. You can then say F, which means file, and you get the menu for file operations.

However, in all cases, it is true that you don't have to specify any data until you get to the bottom level. And that's probably not true here, with this particular application. I think we have a lot to learn from some of those packages.

Monahan: Absolutely. I think that's a good observation. I'd add that one of the best command decoders that I have ever used was on the Control Data 600 series of computer. It was fantastic. If you began typing, the moment you typed what you needed, the rest of the command appeared. It also would not permit you to enter a character that wasn't part of the legal command.

The catch was, of course, that this was on the control console and it used a good deal of the machine cycles to do that. I think that the point just made is that the microcomputers are doing that, because they're dedicating full computing power to the local person.

I do want to make clear that on Figure 1, where I had on the one hand menu and on the other hand command, I don't really regard that as two separate worlds. I think it's a continuum. The description I gave of a menu of having all the options there is the bottom line.

The only justification for not having only menus is that we can't instantaneously supply people with the sort of command--or the user friendliness that we're talking about. We can't instantaneously detect when they're making mistakes and feeling uncomfortable.

Comment: If you're going to provide dual modes, menu and command, you're obviously going to default to menu. That's a double-edged sword, because it gets your users going, but then they tend to want to stay with it rather than move on to the power of command. One way that we try to wean our users is by clearing the screen after a menu sequence has been presented to the user, and displaying the command in the command language, trying to draw his attention to it. And then he gets his results.

So we always give in command language the results of what the menu has added up to. That's been helpful. Not as much as we'd like, because we end up having to have the reference librarian say: "Are you reading the commands?" Interesting enough, only a small percentage actually read them. Most of them just ignore them. But it's a way to move the user along into the command language.

Monahan: What percentage of the users use command language versus menu?

Comment: This comes back to something I said earlier. The default is menu. I think it's 40 to 60 command versus menu. We would like to move it over to command.

Monahan: Any evidence that it's moving?

Comment: Yes, it is moving. With tiny creeps, but it is moving. We would like to experiment with starting out in command mode, but we have to do that on a selective basis. Some reference librarians are willing to let us do that to the terminals in their areas; some aren't.

Monahan: Before you go on, one thing I'd like to relate to that, what about system overhead?

Comment: Command mode is definitely more efficient by a long shot on the system as a whole, especially on telecommunications. It costs money to put those menu screens out there.

Also, while I don't know anything about artificial intelligence, or natural language processing, I'm intrigued with it. We had a programmer that unfortunately left for Wellington and other places, desiring to be an American abroad. Before he left, though, he wrote an amazing SNOBOL4 program that could easily be bolted on the front of the online catalog. We haven't bothered to do it yet, but the way it works is, it just says: "Hi, what are you looking for?" You can type in very, very natural expressions like "I'm

looking for books about airplanes, or I'm interested to know if you have anything on the subject of...." It'll parse away, and it usually gets it right. If it doesn't, if there's some ambiguity, it will come back and ask you questions. In about 99% of the cases, it can come up with the MELVYL command that is the intellectual equivalent of what the user has typed in in natural language. I think that's very promising. I don't understand how it works, but it's very impressive.

Comment: On that, I read an interesting distinction in one of the computing journals about six months ago between English and English-like systems. An English system is something that understands anything that you put in in English. An English-like system is something where, if you put something in and it doesn't understand it, it comes back and says: "Which 'their' did you mean? You mean the city council salaries or the firemen's salaries?" English-like systems seem doable.

Monahan: Yes. I think that the pure English systems will only be doable when computers can understand and replace any of us. But English-like systems, I think, are getting in some sense frighteningly close.

Comment: I don't know how close French-like systems are, though.

Monahan: It's the same thing, it just takes more characters. Are any of the systems represented here using some sort of natural language interface?

Comment: CITE does. I might add that CITE has been installed, and you can all walk in in Bethesda and use it at the National Library of Medicine. It underwent an extensive evaluation face to face with a version of ILS at NLM in Bethesda last summer, early fall. A number of methodologies were used to compare and evaluate them. There's a report that's available, authored by Elliot Siegel, who managed the evaluation of the research project.

But to make one salient point, CITE requires no command language, no learning of syntax and semantics and codes and whatever. It does have some natural language things, and it's also menu driven. It's a mixed situation.

Health professionals, for the most part highly trained and educated, are the users of that system. Staff were also involved in the evaluation. And it was selected over ILS. It's not quite the same issue, because ILS also is a mixed system, and they're integrated. If you know the command language, you can bypass one more step in the menu, and sometimes a lot of steps. So it wasn't one way versus the other. But here's a case where the study clearly shows--and there are others in special library environments, for example--that intelligent, trained expert individuals prefer to use a system--Paper Chase at Beth Israel Hospital is another example--that does not require them to use any kind of command language. And keep that in mind. There are data out there, and there are studies, and they're in place very successfully. The Paper Chase system at Beth Israel Hospital, and CITE at the National Library of

Medicine.

Monahan: That's one of the reasons why I suggested, notwithstanding getting the pure artificial intelligence, that systems like that are now, in practice, becoming available. I think it's important to note that most of those systems are going into special libraries where there is a well-educated, well-defined clientele and, hypothetically, either a small database or, at least, some uniformity among what's being done.

What happens when you put it into a public library? How well does it work in that sort of environment, and how well does it work with some of our other stuff?

Comment: We can't assume that well-trained, intelligent people, scientific and health technicians and so forth, are going to be offended by a lot of menus. They're not, and the data has shown that. They don't search that frequently, and they'll put up with a few menus rather than learn any kind of command language, even a standardized one.

Comment: One thing I do want to say about natural language systems like CITE, the ones I'm familiar with at least are--if you think the concerns expressed during John Schroeder's session were significant, you ought to see what it takes to do some of that.

For example, initially when those tests were brought up, they had to bring virtually everything else down at the National Library of Medicine, because it consumed about 70% of the system for about ten terminals. They've taken some steps to correct some of that, but that's how it works. It is very sophisticated, it has some very powerful features, but it's extremely consuming to the system.

Monahan: I know that I won't get universal agreement on that, but I think anybody who has looked at the LISA has to admit that it's going to radically change their ideas of how we can deliver things to people and what that whole interface does. There's no real trick to LISA except that it delivers a great deal of computing power to the person sitting in front of the terminal. So the real issue becomes cost.

The point is that this is all related to computing power. I think that whenever we talk about the whole concept of codes and languages, then the real issue becomes how much computing power are we going to assume the individual has. That gets back to a dollar issue.

Question: What's the icon for an added entry? (laughter)

Comment: Related to that from the user's perspective, not from the machine perspective, one of the things I've never been able to understand is why systems require so much precision in terms of the command language syntax. I don't understand why you can't build little lists of synonyms in terms of

author--A, AU, the whole word kind of thing, or related kinds of words.

As I use a number of different systems, I get terribly frustrated. I just get confused. I know that I want to do a subject search and can't remember the precise term, and I keep getting these wonderful little error messages. Well, if you can generate the wonderful little error messages about incorrect syntax, by God, you ought to be able to take what I give you and either run it through a list of synonyms or do something in terms of fixing it, and then trundle off to do the processing.

Monahan: We're hitting one problem there. The programmers who, by and large, are allowed to design the command syntax are used to programming languages that absolutely don't allow that sort of thing, so they don't think too much about synonyms. Programming languages have got absolutely nothing to do with the online catalog, but programmers use them all the time. One of the problems is that a lot of programming languages that do that sort of thing for you, programmers hate because of exactly that. So they're probably less willing to grant credit to the user at the other end to allow for that. I think that gets back to the fact that we design some of our command languages as if the user is us.

Comment: I would like to echo the point made about the interest of the professional searcher. We had a situation in training professional staff. Because our system is not menu driven and easy to comprehend, there's a tremendous overhead in the form of an ongoing training program, which is very costly because the retention rate is small. This does not reflect on the folks involved.

Let me just rephrase something. It's been pointed out that an educated user is not offended by menus. My slight rephrasing would be that they're not offended by a system that is comprehensible, as someone rephrased the issue once before.

We're not in a position to throw out what we have developed at some enormous cost, and bring in something perhaps as clear-cut as CITE. We have a slightly different problem, that is, how we take our current command-driven system and make it as comprehensible as possible. We can't initiate a whole series of menus now. We're working from the other end and moving in the direction, we hope, of comprehensibility.

Comment: It seems to me that the software techniques that we're using to parse commands are probably pre-1960. A good way to get the job done is to have the right tools. How many people do use table drive parsers? What kind do they use? I don't think, in general, we're anywhere near the state of the art.

Comment: We can get the language up and running real easy, but the error correction is different. Another order of magnitude.

Monahan: A parser that detects syntax errors may be a long way from knowing what it is that you really want to do.

Comment: When there's a failure, you don't know what part of the production failed. All you know is that it failed. In the table-driven systems you know it failed, but it is not putting you any closer to knowing why.

Comment: You have to be able to invent a certain amount of semantics within the syntax, wouldn't you say?

Comment: I've seen system programmers who weren't out of high school in 1960 still programming with this technology. You either have the interactive approach to computing or you don't. I've had conversations with people in my own company discussing the potential syntax of a new command, and there's the less restrictive way and the more restrictive way. They argue for the more restrictive way because it's easier to detect an error and more clean-cut. My response is when you hit that ambiguity, it's very simple. Just ask the user what he meant.

Comment: Yeah, but you've got to be able to put some semantics in there to do that at that particular point.

Comment: But it was agreed in this case that they could do that. They simply find that this is an abhorrent solution. Asking the user what he meant is abhorrent. They want to program their way out of it, or they don't want to do it. And they have a great voice in these things.

Question: I'm wondering if there's any hard data to suggest that one particular command syntax or another is better. Two character versus four character, for instance? Terseness versus more explicit? How do you handle Boolean operators in the midst of all that? Is there any hard evidence?

Monahan: I'm unaware of any hard evidence. I think that part of the real difficulty there is that the answer becomes so subjective. It's related to how you value a few key strokes over how much it's important for you to know what it is that you're entering or being able to remember.

One technique is coming into common use. We don't have it in our system and I haven't seen it very much at all in online catalogs. That is where there are reasonably verbose words and phrases, but any unique subset of them is legitimate. In other words, you can say: "Copy." Copy works, but so does COP and so does CO, because the system says that's the meaning.

Comment: I read a paper not too long ago that was done by a psychologist who said that five to seven characters is the optimum size for mnemonics or commands, because that's what people were able to remember. If they're smaller or larger, they don't do a good job on it. I don't know how true it is.

Comment: When they study people looking at quantities of items, how many dots there are, they find that same kind of thing. Somewhere around after five to seven, people stop being able to count.

Comment: It's the chunk in short-term memory.

VII. ONLINE USER PROMPTS AND AIDS

Clay Burrows

I was asked to start a provocative discussion and I hope to do that. I will try to keep my comments brief to allow time for extended discussion.

In order to put this talk into context, we should remind ourselves that there are precious few online catalogs in operation at the end of 1983. Considering the thousands of libraries in this country, you must appreciate that only a small percentage of those have online catalogs in operation. More and more are being implemented every month, but the numbers are still very small.

We have only begun to grapple with the problems of how to present the online catalog to the public.

Standards

As we take advantage of the knowledge that we are gaining, we have to be very careful not to paint ourselves into a corner with the systems we are creating. I think the standards discussed here have implications that we may have to face in a few years.

I am concerned that a system that people now consider to be state of the art may turn out to be an albatross in the future. Our role during this time of transition is to develop generalized systems, and not to design the ultimate interface. I am in favor of standards, but I do not think they should be set in place until we learn more about the use of the systems and what type of access is required. I am, however, in favor of working now toward a standard for intersystem communication between unlike systems.

Primitive Operations and Flexibility

Ideally, we will establish a standard set of "primitive" functional operations: direct commands or standardized basic operations. We need to define a set of primitive operations, which can be universally implemented, that will allow for a shopping cart of generalized functions to choose from when designing the public interface. The searching screens built from the primitives would internally generate their own commands based on the system or database being accessed.

During the next several years, we will undoubtedly be going through several transitions as people start conversing with machines in longer and more complex conversations.

We must be prepared for many great advances, but also many failures. Some schemes simply will not be appropriate as our clients become more computer-literate. The students coming out of elementary school will be emerging with more and more computer literacy as each year passes.

We are talking about catalogs and cataloging information, but the systems we are putting together are too expensive to be used just for the libraries' own catalogs. The facilities required for today's online catalogs are generalized search services. The systems could be used for many different kinds of data. The types of access points and search screens are going to change, based on the type of database to be searched.

As we develop functional capabilities, we should be looking toward developing flexible front ends that can be changed, manipulated, and, in some cases, thrown away without changing the basic system structure. One advantage to putting a totally external front end on the system is that search strategies can be developed to access different databases using different options that are not apparent to the terminal operator. The CLR-sponsored Linked Systems Project (LSP) to link WLN, RLIN, and LC employs such a concept.

To search for records with a particular search strategy might require one set of commands for one database and another set of commands to search another database, but the results will be the same. To search one environment or another, the same basic primitives are necessary, so why should the user need to know the difference? With intersystem links, the front end being put on one system could be used for several systems. Using this type of an approach, library users would only have to learn one interface.

To allow for this flexibility, systems must have a utility program that can map the search screen into a set of commands. Those basic primitives can be presented to the library implementors, who can use them to establish very powerful search screens that do not require knowledge of the basic primitive operations.

Providing Multiple Models

The system would allow the library to choose the interfaces that are most appropriate for a particular type of patron and a particular database. A system should be available in several basic starter formats that all follow a logical format: a public version, a research version, and several staff versions, each suited to a particular need. A system for an undergraduate library can be simple enough that a training effort is not necessary. The system can have more than one model of the online catalog to allow the people using the catalog to advance to higher levels of sophistication as their skills improve.

Before the online catalog is brought up, library staff can choose a basic format to begin with and then start working with it, modeling it and

making changes to it, independent of the vendor. They do not need to wait for the vendor to take weeks to deliver a program or months to implement a new display format. These people can work at their own speed. They will suffer some of the frustrations, and perhaps learn to appreciate some of the problems.

Systems are being developed using the personal computer (PC) as a base, to mainframe systems, and everything in the middle. There are many ways to approach this problem with existing tools, but we must be generating generalized systems. We should not be locking ourselves into a given format of a screen or a given format of a menu. We have to allow this facility to be given to the people using the online catalog in day-to-day operations, allowing them to modify it at will, creating the models without the need to call the vendor and say, "Please put the LC card number prompt down here, I really don't want it up there any more," or "Please change my menu prompt, I don't like the terminology used. I would rather use something else."

The key is a set of standardized utilities that allow the screens to be defined easily, allow the error messages to be modified easily, and allow for prompts and text that can be changed as studies are done that determine better ways of saying the same thing.

Who Designs The Online Catalog

We have most of the knowledge to finish the integrated catalog. The technology is here but, clearly, the problem is how to present these capabilities to the public.

We should not think about new design techniques, or how to do it right the next time. We have to work together with the library community and the end users to design and implement the tools necessary to allow for change. Give those tools to the library, and allow them to work together with the users. Together, over time, these groups will implement functionally oriented systems more successfully than a vendor alone can do.

The user community is becoming more responsive, and is more involved in the analysis of interfaces. Committees made up of library staff, scholars, teachers, and students are providing direct input into the design of the catalog interface.

As a vendor, I would just as soon relinquish the responsibility of trying to design the ultimate online catalog, because I am convinced that the vendors, by themselves, are going to fail. We are not the people who use the catalog. We are not the people who have to train users. We are not the people who have to answer the questions that arise from its use. Therefore, systems people should not be the people who design the catalog interface. The programmers and analysts who have been responsible for writing these systems should be cut off from that external interface. Let the programmers and analysts write the code, but let somebody else determine the way to present it

to the public. Let the programmers establish the standard functions, the basic operating systems primitives, but do not allow systems people to write the text that explains the use of those functions to the public.

When we remove this task from the technical staff, we are going to see that the poor input screens, the rotten error messages, and the indecipherable systems will start to clear up as we begin to move these tools out of the hands of the technicians and put them into the hands of the librarians. As each function is described and each error message is built, the systems librarians will have to document the internal procedures and train their staff.

Building Models

When choosing a model to start working with, you must look at the interface to the user. This interface boils down to four processing stages. First, the input of whatever terms are needed for the search. Second, the processing of those terms, hopefully, for a meaningful result. Third, the display of the result. The last step is to indicate to the user what follow-on search options are available. In other words, given any particular type of display from the system, there is a logical subset of system functions that can follow.

The first three steps have been talked about at length. However, more attention needs to be paid to the fourth step, the follow-on search options. For an example of the fourth step, think of the follow-on options available after the display of authority data. It seems logical to have, as an option, the ability to go into the catalog using one of the headings. If you have a call number on the screen, it seems natural to ask for location information. Calling all information up on one screen has performance implications. Screens are not large enough to hold all of the data, and the screens can become very cluttered. What number of follow-on options are reasonable to expect patrons to deal with? There cannot be 50 options. There isn't the space on the screen, and people would not understand.

It seems logical to include in the online catalog different types of searching models that will allow for catalog growth and learning the system. As we design these systems, we need to allow not only for flexibility of one model, but also to have several models: one model for technical processing, another for patrons that never use technical services functions, and others for reference staff that have to bounce back and forth all around.

So as I look at how the systems are used, I can see the need for four or five different presentations of the catalog to the various types of people that are using it.

Economics

Even with the cost of computers going down, the costs of developing systems are still going up. Third and fourth generation programming languages have helped a lot, but trained people are hard to find and cost a lot. The amount of time being spent in design and the number of people involved to help with design makes the staffing problem even greater today, in my opinion, than it was 10 years ago.

We need to integrate some of these other non-library type functions into the existing systems. We are putting in too much hardware and too much sophisticated software to tell the library administrators every three years that if they want more features in the system, they will have to buy another box, and squeeze it in the same computer room, and find another operator to run it, and so on.

These and other financial considerations are forcing us into the arena of designing generalized flexible catalog interfaces to support increases in user sophistication, and to integrate non-traditional databases.

Conclusion

Designing for flexible front ends is a newer concept in library automation that needs more discussion in the literature. Is it a good idea to allow this type of a flexible front end? Is there an alternative that sounds as good? There are drawbacks, and I am certain some of you can identify them. Still, the problem is how to keep our systems flexible enough so that systems designers do not have to run around and redesign, rewrite, and recode every time a new standard comes out, every time a new feature is brought out by a new committee, or a new field goes into a record. If we do not establish generalized and flexible tools, how are we going to deal with future demands?

QUESTIONS AND DISCUSSION

Comment: I don't think librarians are going to have any easier time than systems designers have had in coming up with optimal solutions. We do the sackcloth and ashes bit quite often, and I think some of it is unwarranted. Sure, we make mistakes and do things badly. But I don't think we're quite as dumb about it as some of your remarks might imply.

So I see motivation for generalities to be coming from something else you said, and that has to do with other information applications. My constituency is academic and research libraries. By no means does the library have the stranglehold on information in such institutions. It comes from all kinds of different places, and there's a need to manage information outside the libraries. There's no reason that we should be confining our capabilities to just bibliographic data and other kinds of data in the library. It's got to be integrated into the fabric of information services in an institution. And for that reason I see the need for generalities, not so much that the librarians and others can tune the interface to ever-heightened levels of sophistication.

Secondly, what you imply, I think, is the existence of a generalized database management system with a lot of self-contained language capability, schema and sub-schema definitions, flexible query languages, flexible data presentation languages, and physical data structures that are oriented toward information instead of the formatted data you find in corporate environments. Those are few and far between.

Burrows: My point is that the systems, the languages that we're developing, we don't have to--all we have to do is develop the tools, develop the primitives, develop the pieces that let you get to those various files, let you get to those various displays. But leave the way we present that to the programmers.

Comment: That is a database management system of a highly specialized nature. That's what you're describing.

Burrows: In the database that most of us have, we have the standard access. There is a standard set of access points that are not varying that much. I think that's what we're going to see standardized, more than anything. People are not going to want to walk up and have totally different kinds of access. And we don't; we have author, title, subject, date, language.

Comment: But the art librarian walks in and says, "I have a slide collection that I want to provide some access to. What kind of tools can you give me to do that?" A totally different kind of application. You can let your imagination run wild about what kinds of things like that are out there.

Comment: That's where the dismountable front end really comes into play. That's the way our system works. We have a project going on at Dartmouth that you all have probably heard about. The Dante Society has put up, using the library software, commentaries to 100 verses of Canto Five of Dante's Inferno. They are using the guts of our library software to search this database, using a front end that they wrote themselves.

Burrows: Right. This idea of linking to these other databases or other services implies that the very first screen that people walk up to is going to be different for every institution. One of them is going to allow access to this type of file and this kind of resource directory, where another system is going to show something different. As systems suppliers, I don't think we can tailor every one of those screen maps to handle every one of those cases for every one of those libraries.

Comment: I disagree with you, because we do precisely that. You look at all of our customers, every one of their screens looks very different, because they can have access to any field, anything they want, and they do. An example, an art collection, other special collections. It's mindboggling. Just recently, I demoed one of our customer's databases, which I hadn't seen at all before, and here was a whole different set of access points. They have several different kinds of authors, authors of reports, authors of books and serials, authors of series, all these different access points. That was something different to me. They are targeting it to the specific types of users who use that technical library. The author screens do look radically different.

Comment: I don't understand why you disagree.

Comment: I thought he said it wasn't being done, and couldn't be done.

Burrows. Basically, what I was saying was it has to be done. As we implement access to these other databases, that individual screen has to change.

Comment: The idea that we can afford to have our languages and screens and all these various things different from institution to institution is based on the understanding that we're offering access to different things, which isn't true at the moment.

We're moving into an age where facsimile transmission continues to come down in cost, and where publication on demand doesn't seem awfully far away, where remote access to other people's collections is already orders of magnitude more prevalent than it was a decade ago. As we focus on ourselves less as custodians of institutions, managers of collections, and people who service those collections, and more as people who are midwives or people who assist other people to get to materials wherever they're held and whatever kinds of materials there are, we're going to be faced with a situation where if we don't have uniformity, the user is going to bear the burden of going out to each of these other systems.

People have already made this point in a number of different ways. But I think as we look at the way publishing is going and access to information is going and full text and numerical databases and interactive, if you will, publications--it's a funny word to use because every delivery is different from every other delivery in the world, so to describe it as a publication in the Gutenberg sense is no longer realistic--as we move into these situtions, if we don't have some consistency in these languages, we're going to have absolute chaos.

Burrows: I don't disagree with that. My basic point is that if we put the effort into the right areas today and make sure that our systems are not locked into rigid protocols, as standards are developed, they'll be virtually trivial to implement. We won't have painted ourselves into a corner, and those standards can be implemented, and not even by ourselves, but by the people that have the systems.

Comment: I understand the commercial reasons to allow everybody to specify what they want right now and one can hardly argue with that sort of mandate. But in the long run, the profession itself--I don't think the vendors are going to be able to force it--the profession is going to have to give up that ability to say, "Well, in my library we do it this way." Because if you do that, you're going to be putting your users at a disadvantage in facing the rest of the world.

Comment: Just as I do with the 3x5 card.

Comment: Just as in 1901 when the Library of Congress began distributing 3x5 cards, everyone was forced into that standard. That was not developed that logically.

When MARC was developed in 1968, we hadn't the foggiest idea how it was going to be used. We tried some experiments. You can go back and read about those experiments. They really were pretty pitiful. The amount of data they had to work with was insignificant. I was laughing before when there was a grid up there. I was thinking, "Where does the festschrift indicator fit?" (laughter) Somebody at Stanford, way back when, thought that we had to have a festschrift indicator, and we got a festschrift indicator, and it was premature.

The ARL card distribution came about as a result of the government telling the Library of Congress to distribute copies of all of its catalog cards to all of the hundred ARL libraries in about 1968, I don't remember the precise year. Large numbers of ARL libraries decided then that they couldn't afford their own classification any more. They couldn't have their own heading system any more. Many schools moved from Dewey to LC, because there were going to be LC numbers on the cards.

These are situations where the outside world, in effect, mandated change. I'm not sure we're in a great deal different situation now. I'm not sure that if we mandated standards now that we'd do any worse job than the Library of Congress did in 1901 when it put those cards out and said, "This is what they look like, it's going to cost an awful lot of money to make them look different."

Comment: We have a project going on where we are using the library system to match job openings with job applicants, to match up art work related to the University and bring it up on display when a patron walks into the University so we know we're displaying the art work, for newspaper indexing, and other things. All it is is just a library system applied to different applications, without change.

Burrows: I would assume that the way those people go at that data, they're asking different questions, getting different types of results.

Comment: But it's the same access strategy, just different data.

Comment: I think that for many library systems or information retrieval systems, there are lots of other things that you can use them for. We once seriously contemplated making our circulation systems manage welfare records, because they did the same sort of thing.

I do have some concerns. I'm not against fourth-generation development tools. That's a great idea. I also think that the ability to generate screens is a good idea. The marketplace is probably going to force vendors to make that available. There's nothing wrong there. I think we do a lot of people a grave disservice, though, if the technical people stand up and say, "Well, we really think that the people from the library ought to be designing their own front ends for the common things, the common stuff."

First of all, one thing we say is that the technical people, whose jobs it is to understand online catalogs, who spend all their time thinking about it, are less competent to do it than the librarian who happens to be there. I'm not arguing that the technical people are geniuses and never make mistakes, but we talk about standards and percentage full of screens and things of that nature, then we turn around and try to make amateurs do it.

One of the reasons that a lot of the designs of online catalogs are weak is that the people who designed them, the programmers, didn't know a lot of what they were doing when they started, but they learned. The second, third, and fourth designs of those people are, at least in theory, and in most cases in practice, a lot better. It's a little bit like the doctor who says, "Here's a book on anatomy, and here's a scalpel, go out and operate."

I also get a little bit cynical, and I know this is unfair, but the one thing that programmers hate to do is maintain systems. Frankly, sitting down with users and discussing what they want on their system isn't nearly as much fun as designing the information retrieval system or the job-matching system that is the next tool.

So, yes, we should develop fourth-generation tools, but let's not back away from the issue of maintaining our interfaces. I think we still have to maintain them. I also fight the commercial battle of how many versions I'm going to create, but I think I may be better off doing that than suggesting to my friends that "there's the scalpel."

Burrows: I agree with you that the technical people certainly are learning a lot. They are learning more all the time. Database management techniques are better, access is faster, everything is getting better, but I think we've proven that we are not adequate at providing that end-user interface. I think those primitives that we're talking about, those commands, those basic structures, we're doing a real good job at that, and we're learning and getting better and better all the time.

Comment: I'm not certain that we have any proof that the people out there are any better at doing the user interface than we are. I think what we're talking about is the kind of communication and feedback that has to go on to permit us to be able to do a better job in terms of designing and implementation. You can give screens, and you can give equipment, and any number of us can do that. Equipment and screens, song indices and art slides, and who knows what, all using the same kind of database manager, and all you're doing is providing a different interface.

That's still different, however, from saying that I'm going to be able to come up with a system that's going to interface all of you people here from my one terminal. That's a whole different ballgame, because it doesn't just depend upon my manipulation of the screen. It depends upon the whole structure of your files. There is the place where you're going to have to move mountains, because if we can't even agree that the system being used by Harvard should be replicated across the country, forget all this other kind of interface. You're going to have a major political problem in trying to come up with something that's going to be able to solve the problem of presenting the same kind of screen no matter what database I'm accessing. I'm not saying it shouldn't be done, but talk about a dream.

Burrows: But on the other hand, wouldn't it be easier for this group to decide on a set of standards for commands for operations that nobody is ever going to see, that we know are going to be designed strictly from a functional point of view? We can sit down, I believe, and pound that out. I think we could come up with a set of operations so that, even though the syntax may be a little different from system to system, there would be a transformation that could be applied.

These are the 10 commands in RLIN that it takes to do the job, and here are the 12 commands that it takes in MELVYL, and here are the 17 commands that it takes with your system. Behind these screens are independent interfaces to those different systems. The public can walk up and search different databases using the same screen, behind it the logic to differentiate between the different syntaxes.

Comment: You might be interested in recent developments in the medical library community. The American Association of Medical Colleges commissioned a report on handling academic information. They came up with the striking conclusion that universities are information-dependent organizations. (laughter) And that they're not particularly effective at managing the information.

What is more important to us is that they also came to the conclusion that the library was very well-positioned to develop an institution-wide information management system. They included a lot of blue sky things in that, like personal knowledge bases, and computer literacy, and external database management. Since then the National Library of Medicine has made grants of about $200,000 to libraries, which involve matched federal money about three to one, and they're beginning to develop these information management systems. The Library of Medicine has also earmarked another quarter of a million dollars to be spent in planning information management systems for the next year or two.

I thought it was interesting that these people felt that the library and librarians were well positioned to mamage the information. They didn't say that the proper place was computer sciences, or anything like that.

Comment: Incidentally, if anybody in this room has not yet read that document, you should get a hold of it. Although its focus is in the medical environment, it is as applicable to any academic environment as anything I've seen yet. You'll find it in the October '82 issue of the Journal of Medical Education--Part 2, it's an entire part all by itself. It's well worth looking at. There are people outside of the National Library of Medicine who are clearly interested in the concept for academic campuses, not limited to medicine.

Comment: I was very surprised but gratified to see where our discussion of this subject has led. It reminds me of something that Hildreth tried to turn us on to, with his five charts before, which is that system design is embedded ir a process of organizational design. I'm now speaking solely about the inside of the library, and also its clientele to the degree it can actively be involved.

As Clay said, I believe that the fact that we're all here indicates that public access catalogs may be step 1. Sometimes I think about it as step 50. You know, this is it. This is the final automation project in my life. (laughter) But a smarter view is that it's really just the start of a lot more to come, and that the lot more to come involves a process of organizational design that informs a process of system design.

One of the things that makes Clay's remarks ring particularly true for me is that, as a systems designer, although I think I know it best, I am intellectually uncomfortable with the fact that my clients have no choice. Whether I know it best or not is immaterial, because in the current state of affairs they must work through me for their ideas to become reality.

Intellectually speaking, I'm more satisfied with the idea that I may know what's best. But the tools required to make the system behave are becoming available to people other than the technicians. That's what I regard to be the heart of the fourth-generation language or whatever, nth-generation language. It equalizes power as well as knowledge, so that the knowledge distribution in a group of designers is not skewed toward the technician, necessarily.

I want to conclude these thoughts with a quick comment on de facto standards. Standards writing is a consensus building operation. Some of us spend a lot of time in the American National Standards Committee structure, be it X3 or Z39 or wherever, to create a forum in which all this can be worked out. It is a process not of enlightenment, but of consensus building. A process that can be audited and changed.

I think I first found the phrase "de facto standard" used when I found myself working on a 1401 payroll program, with a 1401 emulator on the IBM 360 something or other. That's when I first started to flinch when someone said de facto standards, because the explanation as to why we needed to emulate 1401 was, "Well, the 1401 is a de facto standard." Does anyone else know what I'm trying to say? (laughter)

That was the first time I started to suspect that "de facto" meant that someone else got there first. So, the reason why I point this out is that de facto standards have their place, but the group of people who judge that they are the standard is not a group that is under public scrutiny to the degree that the laborious process represented by Z39, Z46, and so forth is.

As a person who before he came here read comments 285 through 297 on COBOL 8X, which is the standard X3 is working on right now, I believe that one of the advantages is that the process of standardization is out there exposed. Sunshine is part of it. One of the problems we don't have with the new COBOL standard is people criticizing it for being not innovative. Most of the criticism is that it's too innovative with respect to the installed base of COBOL programs.

Comment: You made the point that some of these tools may equalize power. I think equally valuable, using Charles' model, is that they may equalize responsibility. Within the Library of Congress, our development has been impeded for some time, in terms of the usability of the systems, because the folks who interact with the computers shunned the opportunity to assume responsibility for participating in the development and design at times when doing the design with staff members was quite a feasible option.

That situation has improved dramatically. In fact, there are now folks in our environment who are, in effect, user advocates. They must be involved in the design process because they know more than other folks on the design team who are more technically trained. The balance is shifting in some useful ways. I see your tools as affirmative ways within our institution even now.

Comment: Well, it was 1970 when I was at NLM, wrestling with the problem of whether it was going to be easier to teach the systems designers something about the library, or teach the librarians something about systems design. I'm not sure that the jury isn't still out. Perhaps one of the responsibilities we can take on, as a group, is to teach some librarians that we work with enough about systems design so they can take on the responsibility as you suggest, and they can help us write some of the screen displays and such things.

Comment: I'd like to reinforce that by saying that in the past two years, our librarians--and believe me we have a lot of them in nine campuses and some 100 libraries--they're not beating down the doors to rewrite our screens. That's a reality. I would dearly love, as a provider, to be able to say, "Hey, don't bother me with that, they're your screens." But that's not the way we put the product on the street.

Burrows: But isn't that a matter of priorities? Wouldn't they rather have, as a higher priority, loading of the database and the authority files rather than changing the screen?

Comment: I don't know. There were any number of committees before MELVYL was finally born. They certainly had the opportunity to get that onto a priority list, even if only just to push it down to the bottom. It never got on a list. We have implemented a new feature that I think we're going to learn something from, and it's called the suggest command. The user types in "suggest," and then we essentially turn the system into a word processor and they can type away. We've gotten some really good ideas that way, but we then evaluate them as to whether they're meritorious, implementable, etc.

Comment: Of course, when you were designing it, the committees didn't have a heck of a lot to say about it, because then they were, as we've noted, unaware, completely uneducated, ignorant. As they get more experienced, we find at WLN, they're not so accepting anymore. They would like that change and this change. The sort of approach that Clay has outlined would give them the opportunity, as they learn the deficiencies or better ways to do things, to implement them, in fact.

Comment: So you could have different screens on different terminals pertaining to slightly different situations, depending upon the location of the terminal in the environment. But there would have to be some kind of meta-language in the background. I think that is what you were talking about.

Burrows: That's right. Take any of our systems, break them down, and we have these functional components that we call commands. You can allow people to use a little bit of logic with those commands. In other words, find author such and so, if it has zero hits then automatically do it--go against the authority file with author so and so. And if it's there, if it's a cross-reference, then automatically expand it. Put logic behind where they type their cursor, or what function key they hit, which boils down to more than just one command, but execution of multiple commands that have logic within

them to do some branching. Then we start to be able to implement some of these more esoteric features without having to get into the systems and rewrite and redesign.

Question: Are there any other systems that have a suggest command like this? Are you getting any usable information from it?

Comment: A great deal.

Question: Any examples?

Comment: A lot of times it's matters of terminology that the users don't understand. We implemented a sort command this go-around as part of our list option. Many of the users thought that when we said "order," that they could actually get a copy of whatever. That's typical.

Comment: We don't have a suggest command, but we have a lot of suggest telephones. Some of the suggestions are physiologically impossible. (laughter) Some of them are really useful. We get a lot of valuable information that way.

Comment: I want to go back to standards for a second, because I'm a little worried. We seem to have been talking about standards for the online catalog. Someone used the example of standardization of the car, and I've heard some debate and comments around about whether cars are standardized. I don't believe that's the issue that we have before us. The issue before us is the standardization of transportation. If you standardize the car, if you make that mistake and standardize transportation as the car, then you miss the bicycle, the horse, the airplane, and all the rest of it.

These are educational systems. As they become more adaptive, it becomes less and less possible to generate the kind of macro functionality descriptions that you were talking about, Clay.

I'm concerned about the political power of standards and the level and timing of their introduction. I'd like to see some more discussion at some point during the course of this meeting about that. It worries me greatly when we start talking about interconnection of systems. Then you begin to get a powerful impetus toward fixing, and I think that's a horrible mistake at this point.

Question: Do you offer alternatives? Do we not interconnect? Do we interconnect and not fix?

Comment: Well, I don't know. One thing that's kind of interesting is, presuming that you have the mechanism for the systems to talk to each other, who is going to say what. No one really has talked about that very much.

I'm not sure that I've got anything to say to you. I mean, at a systems level. Or that my users do. We do know from looking at document delivery, interlibrary loan, that the amount of that activity is very low in comparison to the amount of activity that goes on in the local situation. I have a suspicion that that has a great deal more to do with the paradigms under which we operate than it does with what could conceivably be done.

Comment: I think the point is that the interconnection may not be to other library systems. When I'm talking about publication on demand, I'm not necessarily talking about other library systems.

Comment: Well, that scares me even more. If we get into a standardization process in this area now, you can bet that it will become institutionalized within the library community. The whole structure that has dealt with every other standards effort we've ever made will be applied to that effort, and destroy it as far as its interconnection to other kinds of systems.

Burrows: I think standards can be put at a low enough level. We have standards in our network protocol. They don't bother us as systems designers, because we have learned to divorce ourselves from having to worry about that particular aspect. I think we can start to bring that up higher and higher. What I am suggesting is that we standardize up to the level perhaps of all the ways we submit commands and functions to the systems.

Comment: It's basically good sophomore engineering.

Comment: Can I bring an instance up, a different one than we've been talking about? Maybe it will help us in this. We developed back in 1968 and '69, again in that great abstraction we were working in, a series of codings and identifications of diacriticals. I've been told recently there was a German "s" we didn't include. It's very important, I'm told. I don't know enough German to know that, or any German, so it's never bothered me.

But now we're going to be going outside to the PCs, we're going to be going to all these different terminals, people are going to be at home dialing into us, and people will be using a variety of different systems. It seems to me that either the rest of the world adopts the library system, and I don't see any movement in that direction, or we adopt the rest of the world's system and reconstruct all of our various databases to get in sync with the rest of the world in terms of the way they handle other languages, diacriticals and so forth. Either we maintain separate terminals that can only go to library data, or we do some translating, or we recast that. Now, that character set is a de facto standard. Well, it's a real standard within our industry. It looks like we're going to have to give it up in order to go across industries.

Comment: One of the issues that seems to be here is a design issue that has to do with the relationships of form and function. Which is coming first? Are we building a railroad before we know whether we have anything to ship, or are we fixing a railroad before we know whether that's what we want to fix or not? I suspect that we don't need to stop talking about it, we just need to refocus a little bit.

Comment: The talk about standards and flexibility and having front ends--these seem to be sort of counter to each other. I'd like, if I walked from one library to another, or one terminal to another in the extreme case, still to be able to talk to the system. With a whole bunch of different front ends, that becomes a little bit more obscure.

Burrows: It certainly does, and I do not advocate horribly diverse front ends. I think we should be working toward a guideline and toward a standard, but I think it's a little too early. I think if we locked ourselves into a standard now, we'd be cutting off our nose to spite our face. The example I will give there is IBM itself. They decided in whenever it was--1962 or '61 or '58, whenever they designed the 360 architecture--that they would come up with a standard that was going to be good from then on. It has, in fact, been more or less upward compatible all the way down the line. It works. We've got a lot of good software, and it's been working real well. But what if they would have waited until 1975. You always can go a little bit farther and come up with a better standard, and I think we're too early for the online catalog.

Comment: I think what you've proposed is a way to finesse the dilemma of having the diversity while we need it, and then getting around the economic and technical problems and coalescing on to a standard when we understand what that ought to be.

Comment: I disagree. I think we're going to be moving toward a situation where there's increased personalization for people. Somebody told me a story last week about somebody who wrote a program, for their own IBM PC, to go into the catalog and take a look at it. But they didn't like the commands, so they wrote it the way they liked it. I think we're going to see increasing personalization out at that front end.

I think we do have to identify what we need to standardize. We need to standardize the information that we're shipping around. The transportation analogy is excellent. We're not shipping library stuff around, we're shipping information around. John was pointing at that same sort of convergence, and we need some standards on what we're passing around.

In terms of the actual interface, let them go buy whatever interfaces they like and it will turn that command language into, if you will, the machine language. As long as we use the same standardized machine language, we can have a lot of compilers.

Burrows: As long as we have the same gauge on our railroads, who cares what the box car looks like?

Comment: If you layer that properly, then you can pick and choose what you standardize and what you don't. You're not forced into standardizing. You have the opportunity to do it where it's useful.

Comment: Clay just gave the perfect example, because the people building tunnels care how big the box cars are.

Comment: Not how long they are, though.

Comment: If they're on a curve, they do. (laughter)

Comment: This is the most basic kind of software engineering, proper layering.

Comment: I would suggest there are a couple of things that might be looked at almost immediately as far as getting some standardization going. I think that Joe has brought up an excellent possibility here, and that is just naming these various fields, spelling them out, exactly what these are. If we are going to abbreviate, what are the abbreviations? In other words, that's coming at it from the back end. Then you get up clear to the top end, which is what I would call the command structure, there's something like X.25, or asynch ASCII standard protocols, or bisynch, or whatever. We need some standard protocol that says that this is an author search and here's the author's name, and then you and your local system break that down, however you do it in your local system. There is a standard that would be coming across the line from XYZ system to ABC system.

Comment: The application level protocol.

Burrows: I would contend that it will be easier to come up with a standard for the elements of the system that are not externally available, that people don't see, like the internal commands. I think we're going to run into more problems trying to come up with a standard for what you call a title or author, because that's external. People are going to be seeing it every day. People are going to be a lot more rigid about that standard. Whereas command languages, not the command languages you enter at the terminal but these meta-languages as Ed would call them, can be standardized because we do have a fairly good handle on what we've got to work with.

Comment: I think that the two things we absolutely must not lose are creativity and competition. Everyone sitting in this room has designed systems, and is out there to do a better job than anybody else in the room. If any kind of standard gets in the way of either creativity or competition, it's going to be counter-productive.

Comment: Without standards you just cannot do certain things. You can't. You can't have inter-machine communication without some agreement as to how that communication takes place.

But I wanted to comment on this idea of primitive operations. That is a way to facilitate inter-machine communication, because that is the level at which two application processes communicate with each other without having to do it through a front end that is oriented toward a human being.

Comment: There are two different kinds of reasons for the standards we're talking here. One you just pointed to. We're talking about standards for efficiency and capability of doing things. We're also talking about standards purely at the human level, which should be the job of the reference librarians. It's not a systems design function. It's a user specification function. I think they're both important, but sometimes there's some apparent conflict in our language because we haven't separated those two out.

Comment: They come down to the same thing though. One relates to the efficiency of the machine. The other one relates to the efficiency of the people using it.

Comment: Agreed. They're different areas of responsibility as we typically organize ourselves.

Burrows: There are fewer issues in the machine-type things, because you're not talking about humans.

Comment: For machine-to-machine communication, there is the ISO seven-level format. It is a publicly arrived at standard for that very purpose. It lets an IBM, God help us, talk to a Univac. All that does work. More to the point, there are becoming national networks that support those standards, and that are modeled around them. So, at that level, that framework exists.

I fully accept, in terms of the application layer that we're talking about for online catalogs, whether you talk personal level or the dialogue that we're going to have between different types of online catalogs, that we're not ready to do it now. If we do it, we are running a real risk of restricting ourselves. I just ask the question. If we're not ready now, does anybody see even remotely on the horizon, when we're going to be? Can you think of anything in computers during the past 25 years when a language or something like that--the user would say, "Now's the time to standardize."

Question: Notwithstanding all its warts, where would we be without MARC?

Comment: I want to respond quickly to that question by referring back to the de facto standard. Usually, if you don't initiate it before it's ready, a competitive shakeout when market share has stabilized is the trigger. Then it becomes a band-of-thieves kind of standardization. "Let's standardize this, so that the new guys can't get in." This is the difference between user/consumer-driven standards writing and producer-driven standards writing.

This idea of waiting until it's clear what the standards should be--COBOL did not become what it is because it was the best programming language on the table. (laughter) And the 3x5 card did not become what it is, et cetera and so forth. These are indications of producer-driven standards enterprises. I have nothing bad to say about COBOL and all the rest of it, but I do have something to say about it being technically advantageous for people to involve themselves in the formality of standards writing a little

too early, because it takes a long time. If user/consumer-driven standards writing is going to become the norm, whatever we mean by standards, then technically competent people have to get in there early enough so that when the dolts who vote, vote, that the smart guys who are busy doing research don't find themselves with "let's hurry up, let's wire..."

Comment: John Kountz tried to get standardization on circulation bar codes in 1972, when the first CLSI system was out. He wasn't able to do it, and we don't have standardization of circulation bar codes and there's no indication we will ever have it.

More to the point, we've been talking about how we can communicate with each other, and there's a standard that we've totally overlooked. We have been talking as though there was a certain amount of standardization in what we're actually doing. I guess it's a little bit easier to pick up an RLIN search and run it against OCLC than maybe the reverse. In any case, if the cataloging networks we have right now start sending messages to each other--there are different things they are doing, and the indices are probably entrenched enough at this point that there's really no way to get standardization there. So if we work out standardization of messages, it isn't going to help in those systems. Now public catalogs aren't at that point.

Standards have to be set early. Again, the LC card example, the MARC example--yes, they were too early. They were standardized too early. And it's a good thing they were.

Comment: You've got to make a judgment on a case-by-case basis about how crucial the existence of a standard is. There's a lot wrong with MARC, but where would we be without it? RS 232, how would you like to be without that?

Comment: With regard to standards for online catalogs, it seems to me that the whole driving force behind it should be that users on XYZ machines can retrieve data from your machine, and have it meaningful and not in some awful format that they've never seen before and don't know how to interpret or decipher. This labeling thing might be a place to start, because if we were all using those sorts of labels in our own systems, at least it's something, labels they have seen before and that are standard. But in order to connect the machines and transmit a request, some higher-level protocol would have to be established.

It seems like those are the two most critical elements. You handle everything else your way on your system, and then you shift that database to back the other system and they decode it and put it into the formats their users are used to. It seems that that is technically feasible.

The root of the whole thing should be that the users see something that they are used to, and they don't have to go through a bunch of shenanigans that they're not familiar with.

Burrows: I would agree. We have a standard for the data. We can shift records around in a standard. We have MARC, and as a communication format, that's what it's for.

There are levels where we don't have standards. When I create one of these end-user screens, what is the set of basic operations behind that that makes it function? I don't think it has to be standardized, but can we at least document it, so that I know what set of commands to give to RLIN to get the right output, and what commands I need to give to MELVYL to get the right output. We've got a long way to go there yet. LSP is starting to define some sort of inter-system communications standards for commands, but it's still got a long way to go too.

Comment: I'd like to report on the stage that ISO is now at on command language standards, and Pauline Cochrane's involvement with them. We're going to meet again in late fall or winter in London. Pauline has taken a stance, as chair of ANSI Z39, on the standardization of retrieval languages, communications codes, etc. She has taken a stance, at this time, of waiting to see what ISO is doing and working with them.

Someone mentioned the need for a directory of common command names, abbreviations, symbols, etc. There is one that ISO has issued. I have a copy of what came out of their February meeting in Belgium, and a lot of documentation that I was asked to review critically and send back. They want to finalize it at the next meeting in London. Pauline realizes that it would be an ISO thing, that we're not going to be able to vary from it a great deal. We may have to just accept that, with nothing except some additions to it if necessary for our internal retrieval systems and languages.

There's a curious linking between standardization of command languages and command names, abbreviations, symbols. I've seen the proposed list for all of them, exactly what they're supposed to be, what the pound sign is going to do, what the question mark is going to do according to the international standard. It will recommend the abbreviations for all the command names that you can imagine. There's a link between command language standardization and labeled screens. Think about it. Part of the working list are field codes, standard labels, and abbreviations for field codes within the context of command languages, SU for subject--well, I think they recommend ST, at this point. It's getting down to that nitty-gritty. They're about ready to close on it.

One of the things that Pauline wanted me to do was to look at the list of field codes, sort of an adjunct to the command names--abbreviations and symbols that they had recommended. Were there enough? Is it in conformity with what we're seeing in most of the online catalogs? The problem is, of course--why these two standardization efforts link and apparently we have a conflict--if they're going to dictate standardized field code labels, and if we do it for labels in display formats...See the problem?

Burrows: What is their intent, just to standardize the command language?

Comment: That's their charge. They're not going to go any further. It's back in Dublin. That's exactly where they are now.

Burrows: But they're not going to go beyond that?

Comment: No, but there's an obvious, not de facto but real world link between the command language, which is standardized internationally, and labels for title, author, subject, series, and so on.

Comment: Well, more to the point, you've got a basic difference in the way DIALOG and SDC go after data. How do you come up with the unified command language? Is it an umbrella that allows both of those to be subsets, or does it tend to mandate one direction another? Those two firms have used the differences between the command languages as competitive selling points. They are, to some extent, selling the same databases.

Comment: All I'm saying is that if we're going to promote the use of labels for component parts of display formats, and if we're going to promote some standardization of abbreviating those words, then we certainly need to be aware of what ISO is doing, or what Z39 might like to do, on the command language side.

Comment: Don't you standardize function also? Can you standardize terminology without, at least implicitly, doing something about the grouping of tasks that will be performed? One command is slightly different from another command. If you standardize on language, you're going to have to standardize on the functions of the commands.

Comment: They've done this and that's what I'm critical of. In the left column of their final drafts, the functions are there. They are taking DIALOG as their model, and saying this is the way it should be for everybody. They have backed off a bit, but in fact, ISO is doing just that. Somebody said that if you're going to deal with the semantics of the command language, you've got to deal with the syntax. Yes, and the functions too.

There's a two-page piece in the spring or summer issue of Online where Pauline summarizes where they are in relation to ISO. She ends that piece with a set of questions. We're prompted to respond to those question, or add our own questions.

Burrows: I'd like to move the discussion off the standards for this and that, and ask for some feedback on the other half, users aid. What I described is a flexible tool that can create help systems, or command screens, or prompt screens. One of the critical issues becomes what we provide as the training tools, as the help systems for the new generations of software we're coming up with. Whose responsibility should it be to define them? How do we explain this particular error situation? How does help get involved, and how does help get terminated? There are some big problems there. Sometimes you can get into the help system, and it takes you eight days to get out. (laughter)

Comment: In some of the systems, you can't get to the help that you need because there's only one path through that particular set of help screens. If I need a particular help screen, yet the system doesn't allow me to get it, that's a terrible system design constraint.

Comment: You may need a meta-help. (laughter)

Comment: Some of the systems have global and local help. If you put a help in the middle of something, it explains to you what you're doing.

Question: Does anyone have a good feel on how helpful the help screens are? For example, how often does someone move from the state of help screen to the state of end, as opposed to the state of continuing the search?

Comment: We have a lot of data on that. In the last part of the transaction analysis paper, we have a lot of that. If they've logged help, we can tell you exactly what they did after it, two steps after, three steps after, or two before it.

Question: What is the finding? Are help screens helpful? Do people read them? Do they use them?

Comment: You can only tell so much of that. You can only answer some of that, in the transaction log. We're currently doing experiments on the effectiveness, problems, and acceptance of LCS help screens. It's coming slow. I don't know of anybody else that's done any controlled research on it.

Comment: I know we've done some work on it, but I am not up on the findings. I have heard favorable comments about the help screens. There are a lot of them. We believe that they're well written.

We felt it important to keep a simple philosophy in mind. Remind the user of where he's been, tell him where he is now, and finally what his options are to proceed. We also have an automatic help. If the user makes the same mistake three times in a row, we know he's doing wrong and we go through that litany I just gave you.

I have to believe that help screens are useful. Having said that, I must also tell you that I was somewhat taken aback, at least initially, to find out that several of the campuses have organized their own MELVYL training sessions. I thought that was totally antithetical to the design of the system. Yet the reference librarians have very strong feelings about their need to be involved in training their users to use the university catalogs. I don't know what all that means.

Question: Do you know how well attended these sessions are?

Comment: They haven't begun yet.

Question: The follow-on options of the screens that you described--should that be a part of the help screen, or should that be an independent option screen, or be on the same screen that you displayed information on?

Burrows: I would contend it should be an information display along with the data, showing them right there what their follow-on options are, help being another part.

Comment: We've been trying to move something along that direction. We've been writing some extensive help screens, and reached the conclusion that it was fine for explaining some of the possibilities that were available, but that the most likely options ought to be immediately portrayed at the system response levels. Now we're going back and writing new system responses. When we're done with that, we'll go back and rewrite the help screens.

Burrows: I advocated the idea of flexibility, so that libraries can go ahead and create their own searching screens. What about that same concept applied to the help system? Should we turn over to our customers, our librarians, the facility to modify and change the error messages? Should they be able to look at it, and see how it's being used, and in one library have a zero hits message come out one way, and somewhere else have a zero hits message come out somewhere else? The help screens in general, local help and all the rest--how much end-user control should we allow for those functions? How much of that should we be removing from our data processing people?

Comment: It gives Ed's reference librarians something to do. (laughter)

Comment: Surveying that as part of the Carnegie project, the findings were that that's a very, very highly desirable set of functions. They really want to have that capability and control.

Question: Who did you ask?

Comment: The RLG members.

Question: The librarians?

Comment: Yes.

Comment: A couple of the things that our system doesn't do well--our helps are often too long and verbose. Helps should be brief and to the point. You can always call help by typing in question marks, but help ought to be presented by the system when it's needed. When the system can recognize the user's on the wrong track, the system ought to pop up and tell him, "If you're doing a subject search, the best approach is to find one good book and then find all the books that are like that book," and that sort of thing.

Comment: In terms of online user prompts and aids, if the system comes up with no hits, some systems come back with these wonderful little messages like "no record found" or something the equivalent of that, which I think is really atrocious. Other systems insert you into the file so that you can scan or browse the closest alphabetically.

Something like 18 or 20% of the commands entered by the user are reentered exactly as the previous command was. That's a manifestation of "I don't trust it, so I'll try it again" or something to that effect. I think it would be interesting to have a number of the designers talk about their philosophy and their approach to dealing with that problem. What do you do when you get no hits?

Comment: I can tell you what we do. We have a message that's automatically displayed that says, "You retrieved zero hits, for advice and assistance type: help zero hits." Then they type "help zero hits," and they get a full verbose screen, not well edited, that suggests a whole variety of possibilities. It begins by saying, "You might start by not assuming that there's nothing in the database, but look at the following possibilities and then re-execute your search." Now we don't know whether this works better than any other approach, because the design of help screens is a clinical art. Until a few users die on us... (laughter)

Question: When you say, "Consider the following possibilities," do you make those possibilities specific to where the person happens to be in the context of the search and the specific data that he happens to have at that point? Or are they generic?

Comment: The answer is that some of the help screens are very specific in that they refer to what happens at that point. In the course of some of our help screens, we refer people to other help screens, look at "search hints," look at "search conferences," or whatever. So some of them are contextual, and some are just general.

Comment: Joe Matthews has made a point that bears repeating in this context. If you give somebody help, you might tell them why you are giving them help. The reason for the phenomenon of users re-entering the same commands may be that, if they didn't get the commands back, they thought they had mistyped. If you display the command back, and then indicate the error message, at least they can look at what they put in and see how they got that result.

Burrows: When a system gets zero hits, should it try to redivert the same search argument into other files and try to come up with other displays? An example would be that if a system gets zero hits on an author, then give an automatic browse, or if it gets zero hits on a keyword, show the distribution of that keyword throughout the database. Should the system try to take some other course? Or then do we get into the PL/1-type problem, where the system is spending so much time trying to figure out what you did that it's much simpler to go ahead and tell them "zero charlie."

Comment: If somebody enters an author, that's conceptually a browsable index. So we show them the browsable index. But if there are no matches on the keyword, what do you show them?

Burrows: If they go for a title keyword that is not a title keyword, you could show them how that word is used in another context. It may not be a title keyword, but it may be a corporate subject word.

What it boils down to, what John was talking about in his paper, is that if the database is organized correctly, you can quickly retrieve the number of postings given any one type of key, and do that for all the different accesses to the database without a lot of extra overhead. You can actually present a fairly good, almost graphic portrayal of this particular term as it relates to the database.

Comment: My point, though, is this: for some of the things we do, there are incredibly obvious responses to no hits. I think once you start getting into Boolean and keyword, it becomes a little bit harder. Part of the problem with the PL/1 example is that it wasn't so much what the computer did for you, it was trying to figure out what the computer did. (laughter)

Comment: That's why it's very, very dangerous. You're trying to second-guess the user. It's just as frustrating for the user to be presented with something that has absolutely no relevance.

I like Herb White's story about the conference messaging system. I went looking for a message for myself, and I gave my name as DeBuse. It came back, "Sorry, there's no message for DeBuse, but I've got one for Davis and one for Decker." (laughter)

VIII. LINKING OF SYSTEMS

Ray DeBuse

With the current fevered activity in online catalog design and implementation, library automation is clearly entering a new era, one of direct public access. At the same time, however, there are the beginnings of another quantum change in this already rapidly evolving field: we are witnessing the beginning of an era of linked systems. Computer-to-computer communications techniques are advancing rapidly and significant work in the library data processing and information retrieval fields is taking advantage of these advances. This work will result in more efficient use of data that has been entered into machine-readable form, facilitate resource sharing, and increase both the breadth of services and the range of people able to receive them.

A prominent design feature of most recently developed computer-based library systems has been the concept of integration, in which a single database is used for multiple functions. The "fully integrated system" has become the goal, if not always the reality, for most library system designers. Without question the phrase has become a darling of the writers who prepare advertising copy for these systems.

The ability to effectively link with other systems and other databases could allow an expansion of the concept of integration to include its encompassing a multiplicity of systems rationally linked into a functional whole. The result would be a new, much more sophisticated form of automated library networking.

In this brief discussion, I will describe some current linking efforts, explore some of the reasons for linking systems, examine various types of links and their characteristics, and, finally, consider some key elements of the economics and politics of linking.

Links Now or Soon in Place

In recent years, librarians have pressed for the development of two kinds of links:

1. Links between bibliographic utilities and local systems, up to now usually one-way from utilities to circulation systems.

2. Linkage among bibliographic utilities.

In both cases, the major objective is to make multiple uses of a single keying of a data record. In addition, both could allow the creation of more comprehensive union catalogs without the expense of extensive tape loading and redundant storage.

The first of these two demands has led to the various one-way, microcomputer-based facilities intended to allow for the "downloading" of records from a bibliographic database, such as that of OCLC, RLIN, UTLAS, or WLN to specific local systems, such as those of CLSI, DataPhase, Geac, and the like. Innovative Interfaces, TPS Electronics, the Tacoma Public Library, and others have developed such facilities. Typically, a record that is to be captured for local system use is first called up and displayed on a utility-connected terminal. It is then transferred, in display format, through the terminal printer port to the micro, where it is reformatted and queued for loading into the local system.

The second of the two demands has led to the Linked Systems Project (LSP), a joint effort by the Washington Library Network (WLN), the Research Libraries Group (RLG), and the Library of Congress. Funded by the Council on Library Resources, LSP involves the development of an applications level, computer-to-computer link among the three participating systems. By applications level linkage is meant the capacity to exchange transactions at the functional level at which the user interacts with his or her system, or at which data is handled internally by the systems. For example, searches keyed on one system with unsuccessful results on its database may be transmitted over the link and translated into the proper form to allow the same search to be performed on the second system. Or a database update in one system may trigger the same update in a system linked to the first. Using this approach, WLN/RLG/LC Linked Systems will support full MARC name authority record transfer between the participating facilities through three different modes:

1. Search and Response, in which an authorized terminal operator on any one system can forward an online search formulated in the search language of the requestor's system to either of the other two systems, to be searched by the receiving system against its database. The search results are returned in the MARC authorities format. The response is displayed in the screen format of the requestor's system at the requestor's terminal.

2. Record Contribution, in which the staff of selected libraries (participants in the Name Authority Cooperative project of LC) create name authority records on their network system (RLIN or WLN) for subsequent transmission to LC via the link, for inclusion in the LC Name Authority File. Changes to existing records in that file are also transferred in the same manner. Transfers are accomplished in the MARC communications format.

3. Record Distribution, in which new records as well as update records are distributed daily from LC to the other two systems, providing nearly synchronized name authority files at all three sites. The distributed records include those input on RLIN and WLN as well as those created at LC. Again, the MARC format is used during the transfer.

Following implementation of the authorities record exchange, the three participants will implement a similar facility for full catalog records. It is hoped by the participants and CLR that other systems will join in using the LSP protocol once it has been proven. It is intended that the protocol itself be eventually extended to support the transfer of holdings, ILL, ordering, full text, and other kinds of library-related data.

Why Link?

We are now seeing new demands for links, brought about by the burgeoning online catalog movement. Obviously the online catalog must contain catalog records and, in most cases, these records are best obtained from a utility, as is now the case with many circulation systems. But there are different needs as well. Not all online catalogs are based upon the integrated system approach. Thus, in some institutions there is the need to interface the catalog with the circulation and/or serials check-in system to allow ready access to availability status by the catalog users. Some librarians wish to include on-order status as well, requiring a link with the acquisitions system if that is a separate facility. In addition, many of us would like to extend the online catalog to include access to serial literature through subject index databases such as those offered by BRS, Dialog, and SDC. The advancing technology is thus opening new development niche after niche, and there is pressure to fill each as quickly as development resources allow. The user is becoming more demanding, in some cases already asking for access to sources of information beyond what the library traditionally has supplied.

The reasons for developing intersystem links can be generalized. It would seem that there are at least five, not necessarily mutually exclusive of one another:

1. To obtain and update the data necessary to provide a specific service from a particular system. Two examples of responses to this need are described above: the one-way, local system interface with a bibliographic utility and the record distribution function of LSP. The transfer of bibliographic data from a catalog system to an acquisitions system would be another manifestation of this need.

2. To increase the size of the database available through a particular system. The LSP contribution function is an example of a

design feature meeting this need: records input on WLN and RLIN are contributed to the LC database. Links between cataloging systems to support a logical union catalog, either nationwide or regionally, would be another response to this class of need.

3. To add value to the services provided by a system. The "user-friendly" front end to OCLC developed by Case Western Reserve University is an example of a kind of linkage meeting this class of need. Another example is the ability to transfer orders to vendors electronically, rather than on paper. (It should be pointed out, however, that some vendors use current electronic ordering services such as the one offered by WLN only as a form of electronic mail, in which the orders are printed at the vendor's premises and then processed as if they had been sent via the Postal Service.)

4. To allow access to different types of data and services than a particular system provides. The previously mentioned links between the online catalog and other systems such as circulation or acquisitions (assuming they are not elements of an integrated system) reflect such a need. Another example might be the joining of a circulation system and a university's registrar's and business offices to allow for access to ID numbers and addresses of authorized borrowers. So-called gateway access to BRS, etc., would also meet this type of need.

5. To distribute the services of a particular system beyond the limits imposed by that system's resources or political jurisdiction. The distribution of online catalog access is an increasingly expensive activity. Links with regional networks, local area networks, or CATV systems might ease some of the burden.

The linking of systems is not the only way to satisfy each of these needs, but as linking technology matures, it will be seen increasingly as a practical solution. Even though linking is in its infancy now, systems designers are already showing a willingness to abandon the elegance, apparent simplicity, and assumed efficiency of the "unit system." In time, with the adoption of rigorous linking standards such as those now emerging, we may well see the rise of a modular approach to library automation, offering a choice of alternative components for each function, much as we now acquire home stereo systems. Each module might be based upon separate processors, perhaps powerful microcomputers, with some providing database access and others catalog input, the user interface, and so on. Is there the need? As long as librarians demand flexibility and a wide range of differing features there is a need that such an approach might well meet.

Types of Links and Their Characteristics

Needs such as those identified above can be met by links that exhibit either or both of two general functions:

1. The provision of search and response, with temporary display of the response through the system in which the request originated.

2. The transfer of files (or records) from a host system to a target system for retention in the target system.

The first function is illustrated by the facility in which a terminal operator on one system can, by entering a command, have that system dial-up and signon to a peer system. Once connected, the operator can then execute searches against the database of the cooperating system. We see such links now installed between various circulation systems. Such an approach might allow one to access one of the subject-oriented database services or a serials service from one's terminal without having to disconnect from the "home" system. Such "gateways" may vary greatly in their sophistication and simplicity of use. Such search and response functions do not warrant the expense of a sophisticated link, with translations of commands, etc., since the financial benefit would generally be small.

The second function of intersystem links is exhibited by the Linked Systems Project design, which, as we have seen, also supports search and response. With the bibliographic utility/circulation system links now in such widespread use, however, we see a much simpler kind of record transfer than that of LSP, consisting of the making of a local copy of what is displayed as the result of a search on the utility database. As microcomputers become more widely used, so will this kind of "link" with local systems. Many database services are expressing increased fear of the economic consequences of such downloading. A few are taking extreme measures to control the use of the data they hold.

Links serving the record or file transfer function might operate in a variety of ways, including the following:

1. They may operate in real time.

2. They may operate in batch or with the use of queues.

3. Transfer may be initiated as the result of a specific human request.

4. Transfer may be a result of some predetermined condition or set of conditions.

5. If the linked databases are essentially duplicate copies of the same data, they may or may not be synchronized.

6. If synchronized, there must be some kind of linking key, usually a control number.

7. Whole records may be transferred.

8. Only individual fields of records may be transferred.

Despite the apparent benefits of file transfer linkage, most of the links now in use by libraries as well as much of the demand voiced by librarians for linkage focuses upon search and response. In fact, the final result of many search and response links may be data transfer between systems, but without the efficiencies of automatic file transfer software.

Other Characteristics of Links

1. Linkage may be unidirectional or bidirectional. Most gateways are one-way, while the Linked Systems Project linkage is bidirectional, with similar types of data flowing in both directions.

2. Links may be between peer systems or between systems in which one in some way controls the others. Such control may be over link activity, internal system functions, data, or any combination of these.

3. Links may be open or closed. If open, any system that meets the protocol and policy requirements established for the operation of the link may participate. The Canadian iNet is designed as an open network, as is LSP. Both are conceived as webs of systems potentially all in communication with one another.

4. Links may be collaborative, as in the two just cited, or clandestine. Clandestine links would be those feared by many database creators and operators, in which data is copied from a system without the knowledge or approval of the system operator. It is conceivable that someone might wish to develop a clandestine link that would supply data to another system secretively.

5. Links may be message-oriented or connection-oriented. LSP is connection-based, meaning that the transfer of data is established by the creation of an online connection between two systems in which application programs of each system interact with the other. iNet is message-based, in which data is transferred through a "pick-up" function from one system to the other, without direct applications interaction.

6. Links may occur at:

   a) The terminal (Innovative Interfaces, Tacoma's Uniface, etc.).

   b) Communications processors between the terminal and the primary system processors.

   c) The primary system processors in either the teleprocessing software alone or in the applications software as well.

   d) A combination of main processor and dedicated link processor placed between the former and the telecommunications line used for the link.

7. A link may use one of a wide variety of communications protocols, including the following:

   a) Terminal emulation may be the only choice for some applications, although it may present a variety of operational difficulties if used without human interaction.

   b) Proprietary protocols such as IBM SNA.

   c) The International Standards Organization's Open Systems Interchange model (OSI), adopted by the Linked Systems Project and the Irving Project in Colorado.

8. Links may be transparent, semi-transparent, or opaque. Transparency, in which the user does not specify that a transaction is to be directed to another system or database, will almost certainly require considerable translation of commands, error messages, screen displays, etc. Semi-transparent links are those in which the system to which a transaction is to be sent must be specified by the user, but the search or other operation is formulated in the command language of the "home" system, as is the case with LSP. An opaque link would be of the "gateway" variety, in which one must interact with the target system in its command language.

This review of link characteristics has been a cursory one, but it demonstrates the complexity of the subject. There are obviously many faces to the interfacing of systems.

Economic and Political Considerations

Not surprisingly, perhaps, the economic and political factors that surround efforts to link systems may be equally as complex. Economic advantages of a link may include one or more of the following:

1. Avoidance of the cost of purchasing or developing redundant software...if a function is supported by a system with which a link can be established, it may not be necessary to duplicate that function in one's own system.

2. Economy of scale: spread the fixed costs of a specific system over a larger number of users.

3. Avoidance of the cost of redundant storage and/or keying of data. This is the argument most often heard for both linking of systems and for the design of unit, integrated systems.

4. Increased ease of maintenance of individual systems, since they may not need to support as many functions. This is offset, perhaps, by the maintenance requirements of the link itself.

5. Offloading of demand at peak periods in order to maintain performance. This is a theoretical advantage that may not prove real, since linking itself carries a processing overhead.

Costs of links may include:

1. Greater use of increasingly expensive telecommunications channels.

2. The processing overhead mentioned above.

3. Maintenance overhead.

4. More complex database maintenance, particularly for synchronized files. Proper design should minimize this cost, however.

5. The cost of the development of the link itself.

In implementing linked systems, many questions with economic implications are certain to arise. The following are examples:

1. Where does one system begin and another end? In other words, who controls the link?

2. How do you account for resources consumed? This may not always be easy.

3. Must software be developed to handle accounting of royalties on the use of data that is transferred over a link?

4. As systems become more open, with increasing interconnections, how does one avoid unauthorized use and abuse?

Similar issues arise in the politics of linking. Certainly the first of the four questions posed above is as political as it is technical. Other political factors include:

1. Ownership of data, an issue to which many librarians have become highly sensitized.

2. Control of system access.

3. Competitive, commercial factors that may work against linking. Typically, strong, established vendors are not willing to forge links with their weaker competition.

4. The perennial difficulty of arriving at accepted standards.

The last point is an important one. If we are to realize the benefits of linked systems at a cost we can afford, we must agree upon and follow basic standards, for both data and protocols. Clearly, without the MARC communications standard we would not now have a collective store of machine-readable cataloging as large as the one we have created over the past 15 years. Shared cataloging would likely not be the force it is if there were not standard tape formats for the exchange of this data.

Analogous standards must exist for online exchange, particularly if we are to implement an open systems network of linked services and facilities. Commitment and some hard work by representatives of libraries, users, and vendors (both for-profit and nonprofit) will yield a new kind of library networking, the networking of systems, providing the benefits of local control of most library processing but without loss of the contributions made by the joining of libraries in collaborative effort.

QUESTIONS AND DISCUSSION

Question: Why do you feel that if we become involved in linking, we may be moving away from integrated systems? I wasn't sure of the rationale behind that.

DeBuse: What I really said, and I may not have said it well enough, is that we aren't moving away from integrated systems. What we're doing is changing our definition of an integrated system. The present concept of an integrated system may no longer be sufficient. But I'm not arguing against integrated systems.

Question: Which is what?

DeBuse: To accomplish a variety of library functions through a single facility, a single set of interrelated software on a single machine or a very closely linked set of machines accessing a common database.

Comment: There's no reason why that can't continue and extend. One of the things that's happening now is that if you want to use various resources, you end up with 15 different terminals in your library. With a fully integrated system, I would see the capability of sitting at your own terminal and then your integrated system provides you with all the necessary links.

DeBuse: That's exactly what I'm proposing. I'm not talking about a monolithic system.

Question: Isn't the difference a physical one, not a logical one? Logically, you're still trying to do the same thing.

DeBuse: That's a good distinction to make. The reason that I didn't make it is that I began thinking about how we are discussing links generally, and I thought about limiting my considerations to links between systems that reside on physically different hardware. But some of the most exciting work being planned and done right now, in fact, consists in links between quite different sets of software which may well run in the same machine.

Comment: There are really three levels. One is the very tidy link where everything has been designed together, planned together, the old traditional approach: put in an integrated system to control the whole thing. Another is where the level of control the designer has is somewhat less, because you're working out interfaces and bringing the things together.

You've also described the Irving project, which, it seems to me, is a horrendous kludge. It's a kind of layered way of approaching things. We can't get these things to talk to each other, so we'll develop yet another meta-layer at which we take these things that don't want to talk to each other

and trick them into talking to each other. There's a certain amount of that going on. It always struck me that if the Irving project would have taken the money and divided it up among the vendors of the systems they were trying to link, and said, "You only get the money when the systems are linked," it would have been linked a year ago. They'd all be talking to each other.

It seems to me that there's a lot of this after the fact, let's see if we can't glue it together, let's put some mud over the cracks between systems. I don't have a very great prognosis of success for efforts of that sort. I've seen attempts to link systems that didn't want to be linked before, and it strikes me that if people who are making systems are going to resist linking or be noncooperative, then those kinds of projects are really climbing uphill.

DeBuse: That's certainly one of the difficulties that I should have explained further when I mentioned the commercial impediments to linking. That has led to the kind of approach that you've described with Irving, and we'll see a good bit more of that, I'm sure. But my hope is that the ideal solution will eventually prevail. That solution is the adoption of the linking standards by the various vendors, and we will then not have to worry about the mud over the cracks.

Question: Are there linking standards under way other than in the BISAC effort and the Linked Systems Project itself?

DeBuse: Oh, yes.

Comment: The iNet project in Canada works on the OSI problem. Tuesday, we transferred a bibliographic file from the National Library of Canada. They had a little problem in development. It took them most of the summer because of difficulties on their host system, talking X-25--but once they got that down, it was fine.

Comment: There is a case where the people who were talking to iNet are cooperating in the effort, as opposed to Irving.

Comment: Oh, absolutely. If that becomes a success--I don't mean a technical success, but a political success--then it's not difficult to perceive that supporting that standard will become a prerequisite for selling an automated library system in Canada. It will be a kludge for the systems installed ahead of time, maybe, but it will become something designed in the future.

Comment: I don't think you can lay it on the commercial vendors in the Irving project, incidentally, because there is at least one in-house developed system where the systems designers chose not to be a party to a link effort. So I don't think you can lay that all at the doorstep of commercial houses.

Comment: No, I'm not. What I'm saying is that the entire effort seemed to be sort of a shotgun marriage situation.

DeBuse: In my use of the word "commercial," I'm not restricting it to the for-profit vendors. (laughter)

Comment: Or the ones who would like to make a profit, and don't. (laughter)

Comment: I think what troubled me most about the Irving project was that it assumed a very static environment on the part of the commercial vendors in the way the interface was going to be built. I anticipated that the interface on our side was going to evolve and change over time, and I saw this creating immense technological problems--the black box in the center trying to keep up with not just us, but also with all of the other commercial and non-commercial systems attached to it. And I saw it just as a never-ending task that, unless there was an immense amount of funding to keep it going, would never be successful on an ongoing basis. I just couldn't see spending our money as well as the Irving investment in that area. We stated that we really supported the concept of standards development, which would then permit the sort of linkages we're talking about.

DeBuse: This obviously points up the need of standards, and it also underscores the point about maintenance. The system becomes much more complex if it's done in a less than ideal fashion, that is, without standards.

Comment: First of all, in defense of my friends in Colorado, there's nothing wrong with pasting mud over a crack if that's the only way to get the crack bridged. The major difficulty that the Irving project got itself into is that it tried to make transparent not just the communication level link, but also the command and system level link. I don't believe that those two things are the same thing at all. We make a great mistake when we think of linking as a kind of a model that says that, ultimately, anybody who sits down at his or her terminal can do anything that they want to do on anybody's system in the language they wish. That isn't going to happen for reasons that you pointed out this morning. And I'm not even sure that it's a good idea.

The language with which people interact with online systems is a terribly important part of the design of those systems, and a very important mechanism by which the system changes the user. What we're talking about is abandoning that or changing it somehow. What you're doing is taking away one of the single most powerful tools that a system designer has to accomplish the transformations in the user that he wishes.

When we are talking about online, personally oriented systems, we are not serving institutions any more. We are serving direct users. We are talking about educational systems--systems that transform people one way or another--that cause things to happen. We're building tools to do that, and

we're going to get into trouble depending upon the layer at which we approach that.

I would much prefer, for example, to have seen the Irving link proceeding. I think they would have done it a year ago if whoever it was that gave them the money to spend in the first place hadn't given them so much. If they'd simply started by saying that all they wanted to do was to be able to talk to any one of these systems from a terminal line on the other system, the hell with trying to make the command thing transparent. Start there, and then take it from there and see where it goes. We need to give a lot of thought to those kinds of things in each linking project that comes up, as well in the standards that we design.

Ray, I don't disagree with you; standards are necessary. It's just that sometimes they get in the way.

DeBuse: Of course, they're intended to. That's part of their purpose--to get in the way of things that are going to be counter-productive, in fact.

Question: I didn't understand very much of that. Are you speaking against transparency in general, or transparency in the instance of Irving?

Comment: My point, and perhaps I stated it very poorly, is that the problem with the Irving project was that there wasn't a general consensus among all the people who had to effect the thing--a political consensus. I don't think this is the only example. I hear universities around the country saying I'm going to get this system from these folks and this system from these folks, and neither of these folks are terribly interested in talking to each other, but I am going to make them talk to each other.

Comment: That's a smorgasbord approach.

Comment: Yes. As was pointed out earlier, each of the systems evolves. All of the systems are evolving. Somebody who tries to create some rigid link between them without the participation, at some level, of the people who are developing these systems, and then maintain it while each of them is evolving separately, is getting into an environment where, whatever their technical brilliance, the political and economic forces are working against any long- term stability.

It's not clear to me that, when the Irving project started out, it would have been reasonable to expect anybody to foresee all these problems. I'm not trying to whip the people who do it. But the end result is that they've gotten very deep into a situation where the number of variables that have to be controlled is beyond the capability of people to control.

Question: A clarification? What's the status of the project right now?

Comment: They're seeking funds. The specifications have been finished; they added some after a lot of external reviews. They're still looking for an enormous chunk of money.

Comment: You're right that customers and prospects of ours certainly are wanting this unlike system interaction--interfacing. What can we say, as a group? Where are we going when that kind of request or demand is put upon us? What kind of a rational explanation can be given to say when the time will be right, or what events have to happen and what standards have to be developed? What real environment must we create before we can go off and effect these links in a way that isn't just short term?

Because I agree with what you're saying. These systems that are being linked together are only being done that way because of some lack on the part of the systems, not because people really want to link unlike systems. They feel that they have no choice.

So what has to be done before those kind of links can go beyond two or three years, three or four years?

Comment: I would go back to Ray's pointing out the political things, and say that what has to be done is to get agreement of the parties. If you're buying a system, whoever is continuing to evolve that system is now a party to what you're doing. If you can't get that agreement, then you shouldn't do it. Without some level of willingness by the people developing the software to make things easier in links, links are going to fail.

Technical events don't occur in a vacuum. Ray has listed a whole lot of these factors, and my sense is that, in the Irving project, again without history to base themselves on, without previous examples, they moved into a situation where they have discovered that the number of variables is just immense. One of the real benefits of the project, I suppose, is that they've done some very elaborate mapping of what all the problems are. They haven't gotten it out in the literature yet.

Comment: I suggest we could be a little bit more dogmatic on the technical issues. The OSI standard was designed for this purpose; it exists. We're crazy, frankly, if we try to do that again. The work being done in the Linked Systems Project is being done on that basis. That's certainly the foundation behind the iNet trials. I admit I can be accused of being a little bit self-serving when I say that, because we have been supporting a good deal of that standard for a long time. But I think that's really the only way to go technically. So then the political problem is simply one of getting all the major participants, not mentioning OCLC or anybody else by name, (laughter) to agree that we ought to start interfacing, and commit them to a period of time to begin to make that available.

Comment: Yes. It's enlightened self-interest that's going to provide the glue, and if it's not there you can forget it.

Comment: That's really the point. If it's not there, it's unwise to make the attempt clandestinely.

Comment: It won't work.

Comment: You mentioned a lot of costs and benefits in linking systems, but you didn't put any numbers on them. I wonder how much of the linking systems effort under way is driven by true cost benefit analysis, and how much is driven by technological curiosity and things like that.

DeBuse: I would say that it's not an either/or there. Certainly we've not seen, in any case that I know of, a full rigorous cost benefit analysis of linking. We have seen some very strong benefits analyses, and I believe that in the case of LSP, at least, the costs indeed will be justified by those benefits.

Comment: The Council has a modest investment in this. We didn't make that investment on the basis of a cost benefit analysis, because there wasn't anything to look at. There was no reality to measure. Our decision to go forward with funding this effort was based on a conviction that if the capability wasn't there, you'd never be able to produce a cost benefit analysis. By providing a capability and turning it loose on the community, there would be reasonable applications produced to take advantage of that capability.

In essence we're providing risk capital and, in the end, if the link doesn't work, there will have been enough learned to make that worth our investment anyway. We believe at the moment that it's going to work and work well.

Comment: It's like any other piece of technology. There are cost-effective uses of it, and uses that aren't cost effective. You must take it on a case-by-case basis. There are cost effective uses of the link. There's no doubt of that in my mind. But I'm sure there are equally non-cost-effective uses.

It's not trivial to implement. We knew that to start with and we are confirming it with our experience.

DeBuse: It became far less and less trivial as you proceeded, too.

Comment: Well, just getting three organizations together to agree on something like this over a period of two years helped pull the airlines out of their depression.

Question: Ray, can you speak a little bit about the overhead, how complex the software is, some idea of the size of the programs we're talking about?

DeBuse: For LSP specifically?

Question: Yes. Where you are? When you expect to be finished?

DeBuse: We'll be testing the standard interconnection, the lower levels of the protocol, at the end of the year or thereabouts. We should have the actual applications, although perhaps not all applications for all three participants, being tested early in the year.

There are programs being written for the network processors. In fact, there will be two sets of programs. RLG is developing programs for the equipment that they have. The Library of Congress is developing programs for the equipment that both WLN and LC have. That's one level of activity, and that one is nearly complete.

Question: What equipment is that?

DeBuse: WLN and LC are using a Data General Eclipse. RLG is using a DEC PDP 11/24.

The more difficult work is in the application software, where we are implementing the higher level protocols and, in fact, having to modify some of the existing functions of each system to take advantage of the link. There's a good bit of work being done, and it varies significantly from one site to another. In the case of LC and RLG, it involves building a new authorities capability. For WLN, it's extensive revision of what has already existed.

Comment: The difficulty at the applications level has a lot to do with the type of layers, the architecture of the applications. If there's a very clear delineation and a very clear interface between the application and the database management, then searching and retrieval is relatively easy. You have a process analogous to the application process, which essentially transforms the network canonical form for, say, a search request into something that the database manager can understand. The database manager doesn't know that it's coming from outside the environment. But where that layer is not clearly defined, the boundary is not clearly defined, you have structural problems, and perhaps a lot of work to do.

Comment: Let me just describe the iNet thing a little more. The National Library of Canada has the DOBIS system, and they're sort of the major supplier, although the University of Quebec and Carlton University are also seeing themselves as suppliers.

The system takes on the aspects that Ray was talking about where there are queues built up. Basically, people dial in through the interactive

mechanism, look at records, then transfer them to a queue. At some time in the future, the same people initiate a request to the host and say: "I want to get those records now." The host responds, and then the records get transferred in a batch.

About how big the programs are, how tough it is and so on, we didn't do the responder module; we just did the initiator module, so that we could get the group of records that we picked up during the interactive phase. It cost us about $10,000, and the program has about 2,000 lines of code.

Question: From what I'm hearing, there are two or maybe two point n versions of software. What about the rights? How available is it? Where is it? Is it proprietary? Is it in the public domain?

DeBuse: Interesting question.

Comment: I'm interested to know what your answer is. (laughter)

DeBuse: I was afraid of that. (laughter) It's a complex question, because WLN licenses its software. We consider any work done under a grant of this sort not to be salable. That is, we believe that we can distribute it, but that we should not be charging for it. So our LSP software will be supplied to our licensees at no cost.

Now, the development of the software that we are obtaining from the Library of Congress is not supported entirely by the Council. Such software, developed by Federal monies, is in the public domain already.

I cannot speak for how RLG views the software that's being developed there, but RLG does not market its system at this point, so sale of their LSP software may not be an issue.

I believe that this linking software should be as available as it can be to others who wish to use it or to build upon it.

Comment: My guess is that there will probably be several implementations or re-implementations, as is the case with most such systems as this. My question then is: what about the spec? What about the documentation? What is there for programmers to work from?

DeBuse: The project is being very well documented. The documentation will be available. I don't know under what conditions--who will pay for reproduction and all of that. We will certainly make whatever we can available to others who wish to implement the LSP protocol.

Comment: Between the three organizations, the documentation for how you guys are putting links together is going to be extensive. Has there been any thought given to a separate document that does not necessarily talk about

your specific implementations, but allows others to define that extra little interface that needs to be done to link onto the network? Is there a totally separate document that has nothing to do with your particular implementations?

Comment: Essentially, up through session level, draft standards are being employed. They're in the process of becoming standards, ISO session, ISO transport, and so on; there are particular options we've chosen. The documentation of the standard itself provides what is required.

Question: Having linked these utilities, is there going to be a spin-off value in linking local online catalogs to the utilities?

DeBuse: I'm not certain that this initial implementation is one that anyone with a local online catalog wants to run right out and do. But there should be spin-off benefits, particularly the standards deriving from this effort.

Comment: RLG regards the work that's being done on this project as a primary component of its distributive systems activity. We're in the process of initiating discussions with vendors of local systems that our members either have or have expressed interest in, to begin working on linkage with those organizations.

Question: Let me ask a question about the physical links. What have you got in terms of telecommunications media?

DeBuse: Telenet.

Comment: So you just connect and disconnect through Telenet? Okay.

Comment: When Wayne Davison did a piece on LSP about a year ago (ITAL March 1983), he made a special point of trying to list important documents at that point. There's a bibliography at the end of his communication.

I'd like to go back to something said earlier. I think it's awfully important. Someone said that in the Irving project, there were two levels of communication. The one was just getting systems to talk together. The other was the question of transparency. Then we all sat and nodded our heads and said, "Well, yeah, that's a whole lot more difficult to do," and walked away from it.

In the long term, it seems to me that we're going to have to move in the direction of transparency for similar systems. It's reasonable to say that if you're going after this kind of information, you're going to use different approaches than going after that kind of information. But users are going to start getting very impatient about different protocols to get the identical kinds of questions answered by systems. So, inevitably, the one

level of communication implies the other. Again, I know they've been doing a lot of stuff in the Irving Project, and they've mapped out a lot of the problems.

The last comment I'd like to make is that, to the best of my knowledge, common denominators are almost invariably low. As long as we try to work out a common denominator between things that are not inherently consistent, we will be lowering the capabilities. I'm sure this is going on in the Linked Systems Project right now, where what any of the systems will be able to do in terms of searches on the other systems is a subset of what they can do on their own system.

DeBuse: Jim Aagaard has done a lot of work in this since he had a contract to develop the applications level protocol that we're using. And it's precisely what has happened. That's one of the principal problems, if not the principal problem, at that level of protocol. It happens that in our particular situation, with LC, RLIN, and WLN, there is a high degree of similarity in the command structure. But your point is well taken. What is going to happen across the link is indeed less than what can be done on any one system.

Comment: The protocol provides a response that says this form of a search is not supported.

Comment: Well, it's user-friendly. (laughter)

Comment: The solution to the problem is to either identify the primitives into which most of the tasks that someone is likely to want to accomplish on a system can be broken, or identify macros that encompass most of the tasks.

The thing that worries me is even worse, though--or better, depending on how you look at it. As the systems that we invent for particular user communities become more adapted, begin to acquire more recognizable and distinct personalities, the possibility of identifying either macros or primitives should necessarily become more difficult over time, not less difficult over time. Someone mentioned the notion of convergence of systems, and I think it's absolutely wrong. I suspect that more divergence is what we're going to see. Look at the software that's being written for microcomputers as a primary example of that. The stuff is getting more and more imaginative in terms of how users react to it, and what it does to users as they act with it.

Comment: If the primitives can be broken down into lower level components, then systems will have a simpler time dealing with them. The trick is to really come up with a powerful set of elementary primitives, one that is not out of the range of any of our systems to deal with. Even though a combination of 10 of those might not be available as one command on a given

system, the functions of all 10 are available piecemeal. I think that's where the emphasis should go. Then we will be able to simulate the user function, even with systems that don't provide exactly the same types of commands.

Comment: A low denominator is better than no denominator if you need to link to start with. This business of convergence versus divergence--I think that linking is going to provide a vehicle that will assist in convergence. What we're really doing is defining a virtual capability in the application layer, which gets mapped into the real systems at either end. That does have a lowest common denominator effect.

I'll give a concrete example. Hypothetically, suppose there were a linkage between RLIN and a system that used derived search keys. RLIN could implement a search key index to handle those kinds of searches, because we do try to react to the needs of our constituency. I think that linkages would tend to raise that lowest common denominator over time, so that systems would begin to converge on that virtual model that is the application protocol.

Comment: I just want to put before you the question of cost benefit analysis, especially for the utilities. The cost benefit analysis each of them does is within its own context. We're talking about the creation of a larger context, a meta-utility. There is concern about the equitable distribution of costs among the participants in that meta-utility. It's a real consideration.

Comment: That isn't to say, necessarily, that linking utilities is a good idea.

Comment: We all have MARC to database mapping, so if we couched our communication queries in MARC terms, it would be possible to translate them into database terms and we'd be part of the way there. So there is a primitive out there that we can use already.

DeBuse: LSP would never have been proposed if it were not for that.

Comment: One more thing about links. It's important, and has nothing to do with how we do it, or what systems do it or whatever.

Linked systems are going to be slow. The response time is not going to be the same for a search on a local system as it is on one that has to ship its stuff over Telenet or whatever. Our users have to have explained to them that the cost, their time in front of that terminal, is well worth the benefit of getting these distributed systems.

IX. TELECOMMUNICATIONS CONSIDERATIONS FOR ONLINE CATALOGS

Edwin Brownrigg

In my experience with online catalogs, I have found that the worst planning and the worst execution are suffered in the area of telecommunications. There are several reasons for this. First, telecommunications is an area where few people are expert. Second, the technology is changing rapidly, with constantly emerging protocols and methods of transmission.

This paper offers a snapshot of the state of the art of data communications systems for online catalog builders. As such, my snapshot ought to become obsolete soon.

Background

To get a clear picture of the problem, we need to consider two separate aspects of telecommunications. One is transmission medium and the other is protocol. The transmission medium is the physical data path, which can consist of one or a mixture of the following: metal wire, optical fiber, or electromagnetic aether. Protocol refers to the agreed-upon conventions used to move bits from a source to a destination. Of the two aspects, protocols are changing more rapidly than transmission media, although there are some startling events taking place in data radio technology.

Long-Haul Telecommunications Media

There are three general choices for long-haul data paths: the phone company, data network companies, or a private network. The choice one makes will depend on the network's data load, the size and topology of the network, and the protocol to be employed within the network. In a very sophisticated network, diagnostic interaction between the transmission medium and the protocol can also be a factor.

Transcending the issues of transmission medium and protocol is the concept of the architecture of the network. The purpose of a potential network architecture for online catalogs ought to be to make maximum use of available bandwidth while affording enough redundancy to ensure very reliable service. In determining a network's architecture, one must consider issues such as how many bits will be moving, and in what directions; the electronic techniques that will connect the various physical components; the logical protocol by which various parts of the network will communicate with each other; and the geographical layout of terminals, computers, and communications paths (i.e., the network topology) and how the parts of the network will interact.

It should be mentioned here that the load factor in online catalogs is peculiar in that it is unbalanced. Far less data is sent from the terminal to the catalog than is transmitted from the catalog to the terminal. One command can retrieve screens full of search results. This means that, if duplex bandwidth is only available in symmetric quantities, the network will waste around 80% of one side of the bandwidth.

The Phone Company

Most people probably think first of the phone company when the issue of data paths arises. Certainly, if only one or a few data paths are to be installed within a continent, the phone company would be a strong possibility. For shorter-haul telecommunications--say around a campus or within a city--the phone company could also be a reasonable choice. In all cases, phone company lines are generally easily installed. They are expensive, however, and you can never own them. Phone company lines are now available as both analog or digital circuits. Essentially, analog circuits require a modem at the circuit ends, while digital circuits allow direct connection. However, it is not possible to lease other than symmetric bandwidth from the phone company.

In short, phone company lines are useful for small networks with simple topology that do not cover dense distances over 600 miles in diameter.

Data Network Companies

Data network companies are a first good alternative to the phone company. As it turns out, some of them use phone company circuits, and in such cases all of the physical restrictions that apply to phone company leased lines also apply to lines leased from a data network company. However, some data network companies have completely private communication media. Whether this makes any difference to the online catalog system depends on how the data network company prices its service. In general, one can expect their prices to be competitive with the phone company's.

A data network company's protocol may or may not be compatible with the protocol used for the online catalog's teleprocessing. Many data network companies use X.25 as a base protocol. To interface asynchronous ASCII terminals, for example, to such a carrier would probably require packet assembler/disassemblers (PADS), which introduce greater cost and more points of failure into the network. Another example of this type of compatibility problem relates to the bizarre effects that can result when control character sequences are entered into the network. In a user-friendly environment, no character sequences should be offlimits, because sooner or later the network will see them from public users.

Private Carriers

Data communication becomes interesting and risky when encapsulated in a private network. Several choices for transmission are available: terrestrial microwave, geosynchronous satellites, VHF/UHF radio, and combinations of the three.

There are two basic approaches to private networks. One is to "piggyback" on an existing network on a contractual basis. This is not the same as using a data network company--rather it is simply an arrangement for sharing services on someone's private network. For example, an attractive opportunity for online catalog builders is to contract with the state to share bandwidth on the statewide public safety network, used by the highway patrol and similar agencies. These statewide networks, most of which are based on terrestrial microwave links, usually have excess bandwidth available for library automation-scale applications. They tend to be analog networks and are primarily voice-oriented, but there are a growing number of instances in which a library application can be found on such a network. Individual online catalogs, or interconnections between online catalogs, consume relatively small bandwidth and can serve as a source of discretionary revenue to the state agency managing the service.

Another approach is to develop a wholly owned private network. In doing this, one must look at the available technology and its advantages and disadvantages for one's application.

The advantage of terrestrial microwave is that it is a tried and proven network technology. The radio equipment is available off the shelf. Licensing procedures are relatively routine. Bandwidth is plentiful, and it accommodates most protocols.

However, terrestrial microwave has disadvantages that must be considered seriously in terms of inaccessibility of location, reliability of electrical power, and the climate conditions. In order to build a terrestrial microwave system, the paths between network nodes are likely to require repeater stations. Each repeater site involves costs for real estate, remote service and maintenance, remote electrical power distribution, erecting a radio shack, and securing the site.

The major disadvantage of a terrestrial microwave network, aside from the costs associated with its acquisition and operation, is that each repeater station represents a potential point of failure. The reliability factor of the whole network is the product of the reliability of all its points. Over a large terrestrial microwave network, a reliability factor of 98.5% means about 2 hours and 30 minutes of downtime per week. Assuming that one third of that time would be during the early morning hours, then statistically the network could fail for approximately 1 hour and 40 minutes during normal usage hours in any given week.

It is very difficult to convince vendors and service and maintenance staff that the online catalog is a "critical" application and that reliability between 98% and 99% is not acceptable. Moreover, the statistic is deceptive, because it does not include the fact that a service call can take several hours to complete. Add to this the cost of 18-hour service coverage, and the disadvantages of terrestrial microwave networks become yet more apparent.

Geosynchronous satellite technology is becoming increasingly attractive for private networks. As with terrestrial microwave, radio technology for satellite transponders and earth stations is maturing to the point where service and hardware are available off the shelf. Similarly, its licensing procedures are now routine. This technology also supports virtually any protocol at very wide bandwidths.

Compared with terrestrial microwave, geosynchronous satellite communication is more reliable, easier to service, and does not have intracontinental distance problems. One satellite vendor now provides dual transceivers per earth station within a common wave guide and dish, thus providing good redundancy and simple field replacement of discrete radio units.

Because of the asymmetric nature of bandwidth usage by the online catalog (mentioned above), and because of the broadcast nature of the satellite transponder, a simplex satellite radio configuration is an intriguing possibility. Consider a star network consisting of the online catalog computer center surrounded by several terminal cluster nodes. Each terminal node could be transmitting, at a very low bandwidth, user keystrokes over a leased phone company line, even though half of the line would be wasted. The screens full of information being sent from the computer to the terminal nodes could be conveyed over a simplex earth station transmitter up to a satellite transponder, which would then transmit the information simultaneously to each terminal node's receive-only earth station. Only the target node would process the received signal. The cost of each receive-only earth station would be only 10% of the cost of a transceiver.

Future Radio Possibilities

Years of experience combining digital information with radio have demonstrated real limitations but have also given rise to some new possibilities. Since World War II, amplitude-modulated short wave has been used to transmit 5-level BAUDOT code at a very slow speed for international telegrams. Dual diversity systems have provided some redundancy by transmitting the same data simultaneously on two different frequencies so that the receivers could compare and select the best of the two signals for processing. But the overall conclusion is that short wave radio and data do not mix well, mainly because of the erratic propagation characteristics within that part of the electromagnetic spectrum.

With the advent of the U.S. space program, UHF and VHF transmission, combined with frequency modulation, have demonstrated that low power transmitters can communicate over thousands of miles--but only in a near-direct line of sight. The implication for terrestrial data radio is that, in order to achieve a very good signal-to-noise ratio, the limiting factor is distance, meaning the earth's horizon.

Nonetheless, with modulation techniques such as quadrature phase shift keying, very reliable medium-bandwidth communication can be realized between two points not obscured by the earth's horizon. This technique might be applied across a college or university campus where there are clusters of online catalog terminals and local area networks. In such a network of several nodes, redundancy can be maintained by selecting the right transmission protocol and multiple routing, and by rapid hardware replacement supported by cold spare equipment.

Another promising radio application suitable for relatively short distances is discussed later in this paper under "Packet Radio For Short-Haul Communications."

Protocol Issues

I will not discuss here the various types of multiplexing techniques like time-division and statistical multiplexing, nor lower-level electronic and link-level protocols. These are well known and, while they are adaptable to small or simple networks, they have been superseded in the last half decade by new or refined packet-switching protocols more appropriate to the online catalog networks of the 1980s.

Protocols for a packet-switched network for an online catalog must be chosen with future needs in mind. The network should be able to communicate with many different types of networks, and the protocols must thus be flexible enough to allow such interconnection.

A number of packet-switching protocols are in use today, including X.25, originally developed in Europe and widely used there, Systems Network Architecture (SNA), developed by IBM and heavily used for networking IBM and some non-IBM equipment, and TCP/IP, developed by Bolt, Beranek and Newman (BBN) and the Defense Advanced Research Projects Agency (DARPA) for the U.S. Department of Defense ARPANET.

There are major philosophical and conceptual differences among SNA, X.25, and TCP/IP. SNA is a centralized protocol oriented to full-screen terminal or computer-to-computer interactions. It is primarily for IBM equipment. Routing in SNA is established at session connection. This reduces a network's ability to adapt quickly to changes in traffic patterns and to make optimal use of bandwidth, and limits the ability to recover from failures in network links.

For implementing a network, X.25 does not offer a complete set of protocols; it must be supplemented by higher-level protocols, which are not yet widely accepted in the United States. X.25 is also limited in that it is strictly a virtual circuit protocol designed with large public data networks in mind.

At this point, the best approach for online catalogs seems to be to exploit the concept of the internet that has evolved over the last few years. The internet concept assumes that there will be many incompatible networks and that what is needed is a means of concatenating them through gateways.

The internet concept is, at present, best implemented in the U.S. Department of Defense TCP/IP protocols. TCP/IP is really two protocols: Transmission Control Protocol and Internet Protocol. TCP is a reliable end-to-end protocol that goes beyond X.25 to address some considerations dealt with in the higher-level CCITT (Comite Consultative International Telegraphique et Telephonique) protocols. However, it is the Internet Protocol that really distinguishes TCP/IP from X.25 or SNA. While SNA and X.25 are concerned primarily with a single network, IP deals specifically with the problems of network interconnection. Internet addresses include both a network specification and a destination address within a target network. In addition, TCP/IP is a very robust protocol, designed to operate reliably over potentially unreliable transmission paths.

Packet Radio for Short-Haul Communications

One of the most promising new frontiers in network data communications is packet radio. An experimental technology for more than a decade now, packet radio research has contributed to the development of the protocol known as carrier sense multiple access/collision detection (CSMA/CD) which allows a number of devices to communicate over a single set of two common radio frequencies.

The theory of CSMA/CD involves two radio carrier frequencies (f1 and f2), a head or base station, and a number of other nodes connected to each other and to the base station. The base station acts as the head of the network; it could be, for example, an online catalog or its terminal front end. Other nodes on the CSMA/CD network could be online catalog terminals. Statistically, there is a very small likelihood, within a brief sampling interval, that two terminals will transmit (i.e., the users press RETURN) simultaneously. If this occurred, the two transmitters' electromagnetic carrier waves would destroy each other and the base station would receive neither signal. It is much more likely that, during a brief sampling interval, only one terminal will transmit to the base station and receive an acknowledgement of its transmission over the base station's broadcast frequency, f1, which all of the terminals are receiving. In order for the base station to communicate with any particular terminal, an addressing protocol in the form of headers on

data packets assures that only the intended terminal will actually process the message broadcast on f1. All others receiving the message will disregard it because it is not addressed to them.

As small as it is, there is still a chance that two or more terminals will transmit on f2 simultaneously. There are two ways that the CSMA/CD protocol handles the problem. First, each terminal also receives on f2, the frequency that it and all other terminals use for transmission. This allows a terminal, once the user presses RETURN, to interrogate f2 as to whether or not f2 is in use. If f2 is in use, the terminal will wait some milliseconds and then try transmitting again. If the terminal finds that f2 is not in use, it will take a chance and transmit. In the unlikely event that two terminals transmit simultaneously, thereby colliding and failing to receive a "transmission received" acknowledgement from the base station, each will wait some random number of milliseconds and then retry. Here the assumption is that each will wait for different random periods and thereby have little likelihood of colliding again.

Generally, the CSMA/CD protocol has been limited to systems transmitting over coaxial cable. In theory, there are no serious limitations on the protocol imposed by radio transceiving other than bandwidth and the bureaucratic exercise of licensing the carrier frequencies involved. In the case of bandwidth, there is the practical problem of turning a radio transmitter on and off fast enough so as not to introduce latency into the transmission time. As for licensing, there are potentially unexpected problems to overcome, because data, no matter how it is modulated within a radio frequency carrier wave, represents a relatively new kind of radio use.

Assuming that engineering and bureaucratic problems are overcome, there is a broad blue-sky future for radio-based implementations of the CSMA/CD protocol among clustered terminals in an online catalog library environment. The net result could be affordable wireless online terminals.

For example, within a library where it would be desirable to move terminals from place to place, there could be considerable benefit to connecting such terminals to the network not by wire, but rather by the electromagnetic aether. In addition, in older libraries the cost of laying wires or cable in marble structures could be greater than using the packet radio techniques suggested here.

Packet radio can be useful for more than just communications within a single building. There is the notorious "last mile" problem, where there are no resources to connect a device to a nearby network. In a library setting, this problem is typified by satellite or branch libraries that need to be connected to the online catalog, but that cannot justify the modems, multiplexers, and data lines necessary to make such a connection. Where there are several such cases, packet radio may be an attractive and cost-effective solution to the last mile problem.

Terminal Considerations

There is a wide variety of terminals from which to choose for online catalogs. However, experience suggests that the issue is not which terminal to choose, but which terminal types to support in one's online catalog teleprocessing software.

Put another way, it seems more expedient to attempt to anticipate common terminal types that will be finding their way to the online catalog via dial-up connections or local area networks, rather than attempting to find a terminal that is ideally suited to one's online catalog. Of course, one compromise is to choose a special terminal model for dedicated connections, and in addition to permit remote connect/disconnect access from the more common terminal types.

There are several examples of "special terminals" in use today. The IBM 327X and its plug-compatible counterparts such as the Telex with the ALA character set are examples. Also there are some leftover terminal types like the IBM 2741 in either EBCDIC or correspondence mode. In the future there will be the Videotex terminal with multi-layer, presentation-level protocol syntax for graphics and color. Moreover, the device connected to an online catalog may not even be a terminal, but rather another host computer, a network, or a packet assembler/disassembler.

The safe approach to terminal support is to assume a lowest common denominator, ASCII. Any dumb terminal, and personal computers and many word processors, can communicate in stop/start teletype mode. In addition, these are much more likely to be found on a local area network because they are relatively inexpensive and common. Having made this statement, however, it is important to add that there is really no ideal ASCII terminal for the online catalog. It appears that this deficiency is now being overcome by some online catalog vendors who are marketing ASCII terminals that are tailored to their product.

Summary and Conclusions

A telecommunications network represents an enormous, long-term commitment. Although the effective life span of a computer is typically about five to seven years, the operational life span of a major telecommunications network is usually measured in decades. Life spans of this magnitude are necessary to successfully amortize the huge initial costs. Consequently, it is essential to develop a network that will avoid early obsolescence. I believe this can be addressed in part through the use of highly extensible, well-defined protocols.

A telecommunications network for an online catalog can be divided into two parts: long-haul telecommunications, and local networking and distribution. For long-haul telecommunications, the major technological questions are

those of organization--for example, how to effectively exploit the broadcast nature of satellite communications for maximum redundancy at the lowest cost.

Local networks for an online catalog are another matter altogether. Here online catalog builders must beware. Experience suggests that in handling local distribution problems, one should take a "building block" approach, bearing in mind that much of this technology is only now becoming available in the marketplace.

In the long run, the key to success in local distribution seems to be to avoid wiring of any kind wherever possible. This eliminates a whole series of facilities management problems and allows local terminals to be relocated at will. Such capabilities have much in common with tactical military data communications systems.

Many questions still remain unanswered. If you wait for answers to all of them you will never implement your network. The reasonable approach in the face of some uncertainty seems to be to develop an extensible, adaptable network that is hospitable to interconnection with others. You should not have to demand uniformity in order to simplify the job of implementing an online catalog network. The goal should be a network capable of evolving to meet the needs of the future--needs that are unforeseen as well as planned.

QUESTIONS AND DISCUSSION

Comment: First of all, I want to say that Columbia does have a lot of marble, even though it sometimes seems like we've lost it all. (laughter) We still have a lot of marble.

We have a private branch exchange telephone system, better known as Centrex-2. And we also have a private branch computer exchange of the Gandalf variety. Quite wonderful. Now, of course, the modems cost $1,000 at both ends regardless of how you put it. This is just to set the stage for the following question.

The one technology that I'm very intrigued about, other than infrared, is data over voice synchronization.

Brownrigg: Teltone.

Comment: Teltone is a vendor; let's not turn them into Xerox and Frigidaire.

Voice over data synchronization allows us to use the established telephone base to handle data traffic. Not only do you already have the wires, but you also are allowed to continue to use them for voice while you're using them for data. This is a particularly attractive way to reach dormitory rooms, although it still costs a thousand dollars for modems. You don't have to run cables. Since you're drawing diagrams, it seems that the omission was by accident rather than by design. Or do you not like this technology?

Brownrigg: First off, let's explain to everybody what we're talking about. Basically, the telephone company is a worst case situation at the switch. You've got a twisted pair of wires coming out of the switch going into every telephone. However, telephones are more or less ubiquitous these days, certainly in dormitories, at the reference desk, and so on.

You can interpose a box here, and another box over there. What this does is allow you to run data into this box, and then have your terminal sitting near your telephone. You can talk on the telephone and you can also pass data around this line, or near it in reality. They don't interfere with each other. What you're doing is using this wire as a poor man's antenna to convey a radio signal between this point and this point, and then you have an RS232 connection to your terminal. If this distance is less than about a mile, although companies that vend this are much more conservative in their specs, you're in pretty good shape. This will work quite nicely.

The reason why I left it out of my talk is because, as far as I'm concerned, it's just another form of radio telecommunication. I don't think it's that cost effective, especially when contrasted to packet radio. But if you've got a "onesie-twosie" type of situation, and you've already got a

telephone where you need to put a terminal, it is a very reasonable thing to do. It's the onesie-twosie situations that kill you. You've got a cluster of 50 terminals in the library, and that guy a half a mile down the road has got to have one too. That one terminal is very expensive.

Comment: Turning to what you said about long amortization periods for dedicated owned networks, what do you think about the changes that are going to be wrought by deregulation and, in particular, the advent of terrestrial fiber optics, the advent of inexpensive digital communications over those media?

In other words, isn't it kind of a gamble to pop for earth station equipment that's got to be amortized over eight years or more, given these other possibilities? It seems to me the picture is really muddy at this point.

Brownrigg: It's very muddy. A couple of considerations. It is now possible to lease on-premise earth stations for $5,000 a month. These are fully redundant earth stations transceiving two different sets of frequencies. All the bandwidth that you could ever use for $5,000 a month.

The phone company, since deregulation, is becoming a really different animal. In California, I couldn't even talk to them without first going to public bid to get their attention. Now, they come to my office and they've even hired--this is Baby Bell I'm talking about now--they've even hired a consultant to go out and find grant money for our project if we use the phone company as a supplier. They're getting very aggressive.

Moreover, I was involved in a link between Virginia and California, a 56 kilobit link that took Baby Bell less than 30 days to install. That's pretty darn impressive.

So they're getting real aggressive. Now, the man behind the curtain--it's the same people doing the installations. All they've changed is their marketing, as far as I can tell.

Comment: PNAMBC communications.

Brownrigg: Yes. So I don't know how revolutionary AT&T performance is going to be. But it's sure going to be different.

As far as fiber optics is concerned, it's one of these technologies where the unit cost keeps tumbling down. I think there are some technological limitations in terms of repeater stations, points of failure, and so on. But as far as I can tell, in experiments that I've read about, it's pretty darn reliable. What I can't get a handle on at all from AT&T is what real costs are.

I have a particular problem with common carriers, including the telephone company, and that's that my particular links are all intrastate, whereas they're in business for primarily interstate communications. And intrastate tariffs are just outrageous. It's really a problem.

Something that Clay was saying earlier was that our linked systems are going to link slowly. It would be nice to imagine some federal legislation that would do for libraries and data communications what's been done for book rate postage. If we can ship books around at a special rate, why can't we ship bits around at a special rate?

Comment: Eileen Cooke of ALA's Washington Office has been pursuing that. There is some sentiment in Congress for it. The difficulty she has had is in quantifying the amount of communication. If any of you have any information to offer her in this, it would be of very great assistance.

Comment: How about lots and lots. (laughter)

Comment: I think she wanted to get a tad more precise.

Comment: How about, however much there is now, and a little bit more.

Comment: The problem is defining what there is now. She hasn't even made an attempt at projecting library communications.

Comment: I mean in total availability, not what we're using.

Comment: From what I know of fiber optics, it appears that going from point to point, it's very effective. When you have one computer talking to another computer, you get a very broad bandwidth communication going point to point.

The problem in fiber optics has been in trying to split the signal. The signal splitting requires very expensive repeaters, unlike coaxial cable transmission where signal splitting is very inexpensive--all you need is a 50-cent connector and you can split the signal.

If you use coaxial cable connections, you're in a situation where you are splitting the signal many, many times, like for your terminals. It's less expensive, at least now, to do that. On coaxial cable transmission itself, what we heard about was the CSMA/CD protocol, which, in generic terms, is time division in a network. You're dividing up the time among users. There's another way to do it, and that is frequency division. Then you can, of course, take frequency division and time division and combine them together. That fundamentally is the difference between what we call baseband networks and broadband networks. Broadbands give you a lot more communication capability than baseband can give you, because they have both frequency division and time division superimposed on each other.

Brownrigg: In the case of broadband, of course, you're limiting yourself to cable because there's no way you're going to get all those frequency allocations to do it over the electromagnetic ether.

Comment: You're absolutely right on that. Broadband networks have to be coaxial-cable-based. In the telephone networking that you were talking about, connecting to the dorms, one of the major problems in going with a telephone-type network is the fact that all the connections are point-to-point connections, which is fine if you have a hundred or five hundred or even a thousand or two thousand connections. If you're talking about connections on the order of 5,000 or 10,000 or 20,000 connections, then point-to-point network maintenance becomes a terrible problem. A line goes down and it takes you two weeks to find out where the problem was. It becomes very important then to look at local area networks for either baseband or broadband communications.

Finally, I'm intrigued by the possibility of taking your online catalog and putting it on a satellite, because we require supercooling of our computers now. As computers become bigger and bigger, we won't have any cooling problems in space. (laughter)

Brownrigg: I was not suggesting putting the computer in outer space. But that's not a bad idea.

Comment: You only have to start that disk spinning once. (laughter)

Comment: I think it's important to note that earth station technology is moving into the consumer domain. For a downlink that'll operate at either 19.2 or 9.6 kilobits, you can buy a station for about $1,500. I believe that for about $6,500, you can buy one that receives either 19.2 or 9.6 and transmits at 1200 bits per second. These dishes are less than a meter in diameter; I think about two feet in the first case, and about three and a half feet in the second.

They can be sited for literally pennies--you know, clamp them on to a windowsill sort of thing. And you can establish the direction empirically. The biggest problem, of course, is the licensing and, as you say, that is becoming a routine exercise. So I think you're going to see a proliferation of these sorts of things also.

Brownrigg: I should also mention that the kind of technology I'm talking about here is the geosynchronous satellite. I'm led to believe, by discussions with Bolt, Beranek and Newman, that within the next half decade, because of the space shuttle and many other things, we're going to see low-level satellites, just hundreds of them, floating by overhead, each one being a packet switching station. So that with very inexpensive antenna engineering and relatively low power radio, you'll be able to send packets to whichever satellite happens to be within shouting distance of you. Then, depending on where you want your signal to go, your packet will be handed off from

satellite to satellite to satellite until there is one more or less over your destination. Of course, this will all happen in milliseconds. What it's going to mean is that, while the satellite network topology will be ever-changing, each one of those satellite packet switches will be smart enough to know where your source and destination are with respect to its changing topology and hand your packets off. It's going to be relatively inexpensive.

Comment: That finesses the problem with transponder failure.

Brownrigg: Yes, right. They die. They do die. As those of us in the Pacific know, we've just lost our weather satellite. The amazing thing about that was that weather prediction in the Pacific became a real problem, because all of the apparatus that used to be used for weather predicting had been shut down because of this local satellite, which then failed. It was really chaotic for agriculture and shipping because nobody knew what the weather was.

Question: There has been some talk about the cost of the actual earth station, the cost of the lines and whatever. What is the projected cost of sharing time on that satellite? How much is the channel itself?

Brownrigg: That, obviously, depends on who you go to. Gimbel's doesn't tell Macy's. It's very competitive. If you're just sending data, you're not doing broadband analog TV, you immediately limit your place in the market. But if you were to take my situation, which is intrastate California, and put a little bit of scale into the picture, in our case 1600 terminals, it's much, much cheaper than going with the phone company.

Question: What kind of factor?

Brownrigg: Factor? A third.

Question: Is that billed basically on data or on a dedicated line?

Brownrigg: They count every bit going by. The ground station is essentially a capital cost or a lease cost. Then, every bit that goes by ticks that meter.

Comment: Geosynchronous satellites have played havoc with data communications in computers because of distances involved. These low-level ones--what are they, 200 miles or something like that? Would they solve that problem?

Brownrigg: They help. Fortunately, in an online catalog, what's 600 milliseconds? But, you're right. It can be a real problem.

Comment: Propagation delay is not such a terrible problem unless you're trying to poll, in which case it absolutely wipes you out. I did a

back-of-the-envelope calculation to take a look at the terrestrial communication network that we now have and the cost of it vis-a-vis a fully connective satellite network including amortization of earth stations over an eight-year period. If we just replaced what we have today, it would be slightly cheaper. What we'd have, of course, would be full connectivity and extra bandwidth. So, that's the interstate picture.

Brownrigg: You're talking about RLG?

Comment: Yes. So you're not going to save any money. What you're going to do is provide better service, be able to do more.

Brownrigg: And future options.

Comment: And future options, yes. But, it's the eight year amortization. If you amortize over four years, that's a very different thing.

Comment: With the kind of prices you're talking about, it's necessary to stretch it out over eight years.

Comment: Well, the kind of prices I'm talking about are for very limited kinds of things. And it's not clear that that kind of technology, for the moment, would serve us.

Brownrigg: If you want to talk dollars, let me leave you with this. As I said, I went to public bid in 1980. Among others, Pacific Telephone responded. This was before deregulation so it would be an interesting exercise to do again if they'll talk to me. The bid I got in 1980 from Pacific Telephone was a system to support a thousand terminals. They wouldn't bid on 1600 terminals, but they could give me a system to support a thousand terminals for $274,000 a month. (laughter)

Comment: You got two? (laughter)

Brownrigg: Sure. We got a discount on the second one. (laughter)

Comment: You might talk a bit about the incredible bandwidth that's available on the satellite in terms of document delivery--the late-night turnaround in terms of interlibrary loan. One of the real problems with FAX is that, due to the density of the text on the printed page for most interlibrary loan requests as opposed to a business letter, it takes 60 to 90 seconds on the CCITT machine, digital FAX, to transmit that, as opposed to a business letter, which typically takes about 15 seconds.

Of course, those kinds of times really blow away the FAX people, because they're not used to dealing with that kind of text density.

Brownrigg: I'm really tempted to jump into a discussion about electronic publishing, although I won't do it. But I believe that a lot of bits

are going to start moving through networks that relate to the documents themselves, and not just the bibliographic references. There are going to be two varieties: the usual alphanumeric data, but also digital images. When you get into digital images, that's a lot of bits. Let's say, in the worst case, we need super good resolution, so we've got six million bits per page. Do you know how long it takes to send six million bits, even at 56 kilobits? It takes a while. With the type of bandwidth that's available on satellite, we'll be able to overcome that, but it's going to be a slow and expensive process. The cost curve is not going to come down real fast.

Command: You raise an interesting point there. OCLC is proposing to use some of its unused bandwidth on its dedicated network to do precisely that with its link with Information Access Corp.--send actual text.

I work for an organization that wants to talk to some people--that really wants to talk to people very infrequently. Blackwell North America wants to be able to contact its customers every so often and say: "Here are some new books we think you folks ought to think about."

Brownrigg: Is this a broadcast message you're talking about?

Comment: No. I'm talking about a point to point. "Here's a response to claims you sent us early this morning," or "We're ready to receive your week's batch of firm orders," or whatever it would be. We're talking about a small amount of data, yet this data has a fair economic consequence both to the library and ourself.

We also would turn around and go back to book publishers in precisely the same way. Now, we're not in a position to do studies and see if we should put up a satellite, because we're sending a very little bit to a lot of people. From our point of view, it would be awfully nice if there was some sort of neutral network out there that connected all of the libraries, that did not say, "This network is tuned for this application," but rather said, "This network is a network--a neutral network--in which if you want to go to just about anybody for whatever reasons, sending data around according to library protocols, you can do it."

I'm a little discouraged to see that there hasn't been anything of that sort. I wonder if you think it is a reasonable approach, and if you see any movement in that direction. If it's a good idea, is there something any of us can do about it?

Brownrigg: I feel that it's inevitable. What you remind me about is the evolution of the ARPANET. I've watched the ARPANET grow for years now. It used to be that you could go into your network operation center and there was a picture--there was a red line, and then there were a bunch of little spikes coming off of it. This was a computer over in England, and this was one in Belgium, and this was somebody's terminal. Then a few PCs began to show up on the ARPANET. I was there recently, and what I saw this time was a

green line and then a whole bunch of networks connected to ARPANET. So it has become a network of networks, or an internet. I think it's inevitable that there will be a library internet, and vendors and everybody will be on it.

Comment: One thing alluded to is that there is a tremendous element of these costs that is introduced by political or regulatory considerations, and I want to make one observation. In Canada, there is such a coast-to-coast network or internet run by the Association of Phone Companies. It's X25 based, not a library network, but we already have a 1600 terminal network using that. It is the most cost-effective network for that purpose, and it's unbelievably inexpensive. Again, it runs several thousand miles coast to coast, and it's got terminals in tiny towns in Nova Scotia connected to it.

The difficulty is, of course, that there is no such reasonably costed network in this country. The equivalents are, I think, close to an order of magnitude more expensive and don't have the same availability. The one message that comes from that is that a lot of these problems aren't technical. It's the regulations we build. Especially with deregulation now, it's quite possible for two years or three years to just throw all these cost calculations to the birds. That's why I think that the comment about eight-year amortization in telecommunications is really scary.

Comment: This may seem a bit naive, but at B/NA, we're tapping into the Envoy electronic mail system to talk to our Canadian customers, and they don't seem to have any problem with billing us or connecting to us in the U.S. Is there any way the Canadians are going to serve the Americans?

Comment: It just gets expensive across the board. I mean because of regulations. I'm not sure what the source of that is. I think it's got something to do with regulations, but it's also true on across-the-board type private carriers. The network we're talking about in Canada is one that bypasses even the Bell paths; you can go right into X25.

Brownrigg: The problems are not technical, but rather, regulatory. The trick I learned from OCLC is that I can get a 9.6 line from Berkeley to L.A. cheaper by going to Los Alamos and then to L.A.

Comment: I wonder if you can go to Nevada.

Brownrigg: Well, we don't have anything in Nevada. There is no University of California in Nevada--but there is in Los Alamos.

Question: Why doesn't standard commercial electronic mail serve the purpose just cost-eyed?

Comment: Well, I'm talking about sending coded information rather than just free text.

Question: You mean, encrypted?

Comment: Well, we did look at things like, can we set up an arrangement where somebody comes into us looking like a terminal, and we set ourselves up as a Telenet node. My memory is that to be a Telenet node would have cost two, three, four grand a month. It's a piece of change.

Question: Why not just use the Telenet or Tymenet electronic mail system?

Brownrigg: Well, that's my second alternative. It's a data network company, and for a certain scale of traffic that makes sense. But for very large-scale applications, it's out of reach.

Comment: The crossover, the last time we calculated it, was eight hours of use a month, given the typical terminal traffic. At that point, it's cheaper to go with a dedicated line.

Not to reduce the level of the discussion, but one of the interesting things that could happen immediately is just to permit dial-up. Very often, there's a lack of sharing, because simple dial-up is not encouraged or permitted. Again, it would seem to me that starting somewhere is better than not having anything at all.

Brownrigg: I'd be interested in hearing other people's opinions on this. We, of course, have had dial-up access to MELVYL for a couple of years. But the dial-up accounts that we have for so-called friends of DLA, people in this room for example, exist because I want to know what they think about our catalog. So I go through the whole administrative rigmarole of setting up an account for such friends. We've got passwords and the whole thing, because I don't want just anybody seizing my lines and accessing my computer.

What happens when we put large concentrators on each of our campuses and some of them are connected to answer modems so anybody can dial-up? The way we will configure those lines, when you dial-up, you're on the catalog. You don't have to give a password, nothing, because that's user-friendly public access. Well, what is public access? Does public access mean that an information broker can perch on our catalog all day? Dial-up is a tricky issue, and I don't know what the answer is.

Comment: While we talk about accessing systems and linking and sharing, the use of computer time is still the use of computer time, whether it's going to be on some very sophisticated communication system or somebody dials up and uses the darn thing. Some kind of sharing could be taking place now if that simple function were permitted. In many cases, we're talking about development and use of sophisticated communications, when we haven't even agreed that we're going to allow anybody to dial-up to the system.

Comment: The problem we cited is beginning to be addressed by people who make switches and concentrators, by putting protection at that level.

Brownrigg: Right. That was exactly the answer I got from one of my campuses, not to worry because we have this password sequence. So I, of course, asked for an example of the password, and it's a very ugly, user-hostile sequence that the user has to go through before he can get on the system.

Comment: But if they can dial-up...

Brownrigg: That's right. If they go to the trouble to dial seven or nine digits, that's pretty unfriendly. They can type some signs and slashes.

Comment: Our attitude, so far, has been that anyone can come in and stand at the card catalog all day. If we're going to eliminate the card catalog, we have to provide the same capability without charge. It hasn't cost us that much yet. Maybe the attitude will change.

Comment: They have to come into your library now. It's a little bit different when you can dial it.

Comment: The problem we've been grappling with is how many dial-ups we should have. What's the ratio we should have of dedicated versus dial-ups? Does anybody have any feel for that? I have a feeling that it's a dynamic thing. If we have a figure today, it's not the figure tomorrow.

Brownrigg: That's right. I think the ratio is going to get more and more top heavy in favor of virtual connections. I think we're going to have relatively fewer and fewer dedicated terminals.

Question: Do you have any hard data on ratios?

Brownrigg: Yes. Right now, we're configuring fifty/fifty. It's expensive to put in dedicated terminals. The joke is that we installed about a hundred MELVYL terminals around the state of California, putting them all essentially next to the card catalog, but the professors still wouldn't walk over to use them. They'd send their graduate assistants. Then I started watching the requests of dial-up accounts. The professors with their PCs log-on from their office and use the system directly.

Question: Are you providing essentially 1200, or 300?

Brownrigg: The dial-up is at both 1200 and 300 bps. The 1200 lines are still a little expensive--$800,000. If you're an occasional online catalog user, 300 is okay.

Comment: There is a formula used by Tymshare people for the number of people that could be using the system with one dial-up, two, three, four. It is interesting that with one, you can get 8 people, but with two, it turns out

to be 64. I don't remember it precisely, but some of the Tymshare people have those kinds of figures.

Comment: We've got a local area network in Guelph--Gandalf.

Brownrigg: Is that really a local area network, or is it just a big switch?

Comment: It's just a big switch. It's relatively cheap, and it does the job for terminals pretty well. We've got 800 terminals that have these modems on them. We have eight of those modems in our library computer. We only get queuing on those things about 10 percent of the time. The switch has the capability of telling when the lines are queued and so on.

Question: Do you have any idea of the extent of the queues?

Comment: It's usually never more than two.

Question: Is 80% availability of a network as bad as 80% availability of a machine? Is that totally inadequate? Does that mean that enough people aren't getting in that it could cause a problem?

Comment: We're not getting any complaints. We could put more ports in if we needed to, but we're not getting any complaints.

Comment: On the availability issue, our experience has been that when ports become unavailable, then people who get on don't get off. It creates a duplicate problem of unavailability rising, so we'll probably work on the idea that the queue line should never exceed two. Which means that our terminal-to-port ratio varies between two and three. This is while we have 2100 terminals on campus. Now that we are going to something like 4500 terminals on campus next year, we are going to change our ratio to a one-to-five port-to-terminal ratio. And we hope that will work.

Brownrigg: We use a technique that Columbia University used for years. Maybe it's now discontinued, I don't know. An express dial-up line--you dial up, you get the line for whatever the privilege is, 5 minutes, 10 minutes, and then it cuts you off.

Comment: What we'd like to do on our terminal is that if there's no traffic for x number of minutes, it just cuts them off.

Comment: I want to make a comment about theft of service. I feel very exposed, in New York City, to the free-lance bibliographers who've always been somewhat of a problem in terms of our journal and monographic collections, especially during the week. We don't check people at the door when they come in, so the private entrepreneur has always been a bit of a problem to a library like Columbia.

The way we're approaching this is that the catalogs on campus that are part of the Gandalf Network are user-friendly. If a person connects to the port that the catalog is available on, they don't log on. We draw the statistics on the basis of the port class. However, a person who dials in via the telephone network connects first to an IBM VM machine and then passes through to the library catalog, so they have to log in to the base operating system and then call the catalog. This is how we get the password protection outside the library application.

Question: You don't charge for that?

Comment: No. Actually it does, but then we credit it. A user accumulates a charge, but they're credited because the charge-back system tracks by application.

If they have the equivalent of a fee borrower card, however, then the credit is not applied to the account. An invoice goes out for all the VM charges. So they could be a general fee-paying VM user at Columbia University and their item charge, for instance, is seven hours of connect with the library catalog. But if you're a funded research user, you get all your itemized VM charges, and then there'll be the itemized library catalog connect charge with an applied credit. This is very rudimentary charge-back. But there is not only the protection, there's also the fact that we're going to want to charge some of our users on campus. The engineering faculty member, who's honest enough to tell us that he's using the computer resource for private consulting, is going to pay for access to the library catalog on that account. We don't deny him access; we try to put together a charge-back procedure. Of course, it requires self-disclosure. (laughter)

Comment: Of course, everybody's honest.

Comment: They're obligated by their contract to the university to disclose use of university resources for this purpose. And some of them do. (laughter)

Question: I'd like to come back to the last mile problem. You quoted the figure for California of $274,000 a month. Would that include solving your last mile problem? And you're obviously not going to take that solution; what are you going to do in the next two to five years for the last mile?

Brownrigg: Well, I much prefer to talk about the long haul. I have taken the position that because I can't control the last mile problem, I'm not going to be responsible for it. My responsibility will end where packet switch devices are. Going on from there, it's going to be up to some kind of negotiation between the university librarian, the computer center director, and the telecommunications officer. I've got nine different cases, and they're all different.

Question: They're going to fund that on campus, then?

Brownrigg: Yes, they are. That was one of the things that we learned with our prototype. When you're doing a prototype, that embraces a whole host of sins; one of the things I did was the last mile, and we lost our shirt in every case. That was just on eight terminals per campus.

Comment: You might mention that some of the campuses are issuing RFPs for local area networks, and some are looking at directional radio, and just a whole variety of things.

Brownrigg: The point is that we had to choose a network protocol that had the internet protocol, because you had better believe that the University of California with its nine campuses was eventually going to be looking at nine different local area network protocols.

Comment: On the local area network for campus, we just went through that process. We plan to connect 66 buildings, every dorm and every building, to the central facility. Half of it is already done. The cabling plan is going to cost just under $250,000. Each connection will have a modem cost of $900 at both ends, together $900, and a one-time installation cost of $75, which will bring the LAN to the exact location we want it. Compared to the telephone connections this approach is about a third of the cost.

Brownrigg: I'm glad somebody mentioned cable, because there is an interesting example at the University of California of how you can get into a creative relationship with your local cable TV folks. The cable TV company in the Riverside area wanted right-of-way to put cable into the dormitories so as to sell service to the students.

The University said, "Yes, we'll give you the right-of-way under the condition that you bring the cable in and terminate it at every building on the campus." And for the cable company, this was no problem. They ran cables to every building, terminated them, and then laced the dormitories with the cable. Now, the University, for free, has had the basic backbone of a local area network put in.

We've also been talking with the cable company in Davis. We have a clear shot off of Sutro Tower in San Francisco, which is a tall, monstrous structure that has a radio antenna on it. If we can get right-of-way, then we will be sending our screens out on channel 82 to Davis, to the local TV cable company. We'll be able to get the screens out that way, but it's only one-way. We'll still have to bring the keystrokes from the terminals back in over a 9.6 or 19.2 line.

Comment: Rich Sweeney of Columbus has some experience using a cable company, and he says that they encountered two problems. Number one, they don't have redundant, fail-safe electrical supply, and so every time there is lightning strike within 100 miles of Columbus, the cable system goes down.

The second problem was that the channel and linkages, which the library was using to communicate between the branches and the CPU, has the lowest of priorities, since it has nothing to do with programming.

Brownrigg: I told them that. We get the same treatment on the California state microwave system. We have a link, an experimental link down in Santa Barbara--and it's wonderful, it's 56 kilohertz. But when you get atmospheric inversions in the Valley, the signal just goes away. Or when it rains very hard, the signal goes away. Or when traffic gets jammed up on the network, the library is lowest priority. There are political as well as technological reasons why state systems aren't all they could be.

Question: Michael, could you talk a little bit about NYU and the installation of a local area network there?

Comment: Well, they have already gone onto an online catalog, and it's being extended now. NYU is trying to address a number of requirements. First of all, the same requirements we're all talking about. They're also looking at the fact that they're going to run not just the online catalog, but also the office communication systems and other university systems. There are already terminals that go out to university computing in the library, so the purpose of a lot of this is to link up the systems. They will probably link multiple local area networks together.

NYU has another problem, related to something that we have been talking about. As a collection of semi-independent institutions, they'd like to bill people who use their online catalog. And they've had a problem, which is that unless people identify themselves, and that goes into privacy again... We originally discussed the concept of a code with them. You know, somebody just says, "I'm a code four user." The difficulty with that is that the codes become, in essence, public knowledge, because there are thousands of people who use them. You don't get any useful information because people change jurisdiction. It's a real problem. Nobody wants to have to say, "At the library, we promise we're only going to use that code to bill you, we're not going to track what you've actually looked at."

Comment: I wonder if this hesitancy about user identification will go away. People are going to use the online catalog, not only to find whether something exists in the collection, but also to ask for it. I think the hesitancy is a temporary phenomenon.

Comment: Let me share one thing that we have done in Buffalo, for Buffalo-Erie County Public Library. There is a central data processing department there, and when Buffalo-Erie County wanted to install a network, they had to go through the county data processing department. The advantage was that when they began to work out telephone lines, what they did was to design a communication network that included the Police Department, all the various places out in the county, the county assessor's office--and the library.

What resulted is this communication network with various nodes of multi-drop lines, all sharing the cost of communication. I don't know how many other places could do that, but that is another option, at least for a small local network, to begin to solve some of their problems--just flat-out share the cost of the line.

We also do that between Buffalo and Cleveland, and we'll be going up to Rochester and Cedar Rapids with dedicated lines. They're simply going to share the cost of those lines using, in this case, DECNET. But again, just sharing the cost of lines.

Brownrigg: You have one thing going for you that should not be taken for granted, and that's that you've got New York telephones. I've had nothing but the best of experiences in data communication with the New York telephone company. That's why I was so shocked when I got to California.

Question: Can I ask a less detailed question? A real fuzzy question. In your system right now, the online catalog system, what percentage of your total cost, however you're billing and to whomever, are telecommunications? And do you see that percentage changing over the next 5 and 10 years? Do you have a sense of the direction it will change?

Brownrigg: Up until now, it has been a relatively small cost, as we only have a hundred terminals. From about now on, as we order three earth stations, one for Berkeley, one for L.A., and one for San Diego, and also as we purchase boatloads of packet switching gear...The cost is about 50% of my equipment and operating budget. I expect that in the future it will be the single largest cost of my operation.

Comment: I think it's about 15-20% of the expense budget at RLIN now. However, that includes communications front end equipment; it includes amortization of equipment, as well as circuits.

Brownrigg: You're talking personnel and everything else?

Comment: Yes. I'm talking about personnel from the network control center, all of the control equipment, all of the front ends.

Brownrigg: To get back to the point I made earlier, you have to bear in mind that my case is particularly exciting. because when the state funded this project, telecommunications wasn't even built into it.

Comment: We're a smaller user, we run a network of serials control that is thinly distributed over a wide geographical area. We have about 25 library users. Telecommunications is a very, very large cost of the total network. It's something like 75-80%. It's not only the cost, but when you're a smaller vendor, you can't hire a telecommunications expert to sit in one office and handle all of that sort of thing, as some of the larger networks

do. It's a real big administrative hassle. It's a very volatile area, so that any time you want to know how much it costs to connect to a library in Houston, you've got to dial up nine different vendors, and figure out their charging algorithms. It really is a big pain in the neck.

Question: When you say 75%, what is the base you're using? Are you including personnel salaries and the whole works?

Comment: Yes, the whole works. It's a big cost. So to get around it, we've done a couple of things.

First of all, you've got to be innovative and figure out different ways to use the telecommunication network to reduce the unit cost. You've got to have a network to run series control. You're also going to do some interlibrary loan, and you do message switching--things that you might not otherwise do if you didn't have a whole telecommunications network.

The second thing we've done is use distributed processing. You can put a lot of the work on micros in the library, get them online to serials control, but offline from the national net.

Comment: I've been looking at telecommunications costs as a percentage of our total automation budget. As nearly as I can tell, with 25 terminals four or five years ago, it was about 2 or 3% of our total automation budget. Now with terminals inching toward 100, it's at least 15% of our automation budget before installing our online catalog. I hesitate to think of what it's going to look like next year.

Comment: I've seen it time and time again. When people start costing the installation of the local catalog, they ignore the operational aspects of telecommunications. It is a big hassle to run a network with 100 or more terminals.

Comment: It's not the least bit of fun, I tell you. It's the worst kind of work.

Brownrigg: Finding the people with the right attitude at the right price to put up with that kind of "fun" is also hard.

X. SUMMARY

Brian Aveney

I'm reminded in this situation of a wonderful line from Fred Kilgour at the 1976 ALA preconference when he announced that he'd be brief because he understood he stood between the audience and the bar.

I also want to observe that in the present state of the art, the online catalog is like the old joke about the piano-playing dog. The wonder is not that the dog plays well, but that the dog plays at all. I think we have been trying to get up some extremely complex systems. We've heard enough in our discussions to understand the complexity of bringing up some of the systems that have been brought up, and hopefully we're starting to move into stages where we can go through iterations of design improvement.

At the same time, we're finding that the things we're being asked to do are changing constantly.

I'm basically just going to make isolated comments about some things that were said. One is on standards. I want to repeat that the LC card format is not a standard that was brought about because people thought it was time to standardize, but became a standard because of an economic act. LC started distributing cards, and everyone started using the LC card format. The MARC format was done the same way--it was a premature standard.

The Library of Congress Subject Headings, which we use, were a premature standard. The field wasn't sure what subject headings should look like but, sometime in the 1890s, somebody published a book. The original book was the American Library Association Subject Headings, but it very quickly became the Library of Congress Subject Headings as they took over the authority of producing it. What happened is simply that the existence of this thing resulted in people using it.

I indicated before that the distribution of the depository sets back in the late 1960s, when the 100 ARL libraries got copies of all LC cards, was what finally turned the trick for a lot of ARL libraries in moving to LC classification, both from unique classifications and from the Dewey classification.

After we talked a bit about these standards, Jim Corey said to me: "Well, what's going on in acquisitions?" And I think it's a useful case in point, the BISAC effort. The BISAC effort started out by reaching a political consensus that there must be a standard and by getting the people talking to each other.

One of the things that we did, that I think has been a big part of the impetus in the BISAC effort, is that my company (B/NA) decided to hold a buffet lunch, a closed session with invited people. We invited all our competitors, all the networks, and all the turnkey vendors. The following ALA, Baker & Taylor did the identical thing. Midwest Library Services did it last time, and I believe OCLC is going to do it at the next convention.

We really haven't had big agendas and talked business. What we did do was to get everybody into the room, and get a general agreement that it was in everybody's economic interest to have a standard. From then on, it became much less important to most of the business people in the room what the standard was. But they were all agreed it was in their economic interest to have a standard.

And I think there is a consensus--a sense at least, among many of us, that some standards would be useful in the area of online catalogs. What we're concerned about is that we don't standardize things, and then have them trip us up shortly thereafter.

Someone commented that we tend to talk in the speed of our discussions with a lot of either/ors. And I think standards are one of those areas where the either/or doesn't serve us well. We can arrive at certain kinds of standards and a statement about general kinds of standards. My own sense is that labeling screens, as we get bigger and wider audiences, is going to be the sort of thing that we'll look at as a kind of standard, and whether the exact labels are worked out to everyone's agreement--the important thing is we have a general sense that, when someone goes up to a terminal, they can comprehend what they have. This is going to be much more important as we start spreading online catalogs across the land.

A brief comment on costs. We've talked about the expense of a lot of our systems. One of the things we haven't talked about here that many of us are aware of is, for example, the fully loaded cost of a reference library. You've got to include salaries and health insurance and chairs and lighting and all of those various things--that's every year. We start looking at the cost of "expensive" front ends, and I think it puts those expenses in a little bit more perspective. We haven't been talking about some of those kinds of tradeoffs. Michael Gorman observed, during a preconference at ALA in June, that the history of library science is the history of replacement of human beings by systems.

Now the systems may have been something like the Dewey system, which hasn't got anything mechanical about it. It's a virtual system if you will. Nonetheless, it replaced the fellow whom you went to in the German library and said, "Do you have any good books on so and so?" and the fellow goes over and says, "Well, this is a good work, and this is a good work, and this is a good work." So the catalog itself, when it was invented, was something that replaced human effort. This is the entire history of library science, and we're just marching down the same path.

Just another brief, extraneous comment on the existing literature. Someone got up during one of the sessions of the LITA Conference and berated speakers on an online catalog session for having ignored all the things that were said 15 years ago in a previous institute. We have been told to look at previous literature.

I don't think that either that person's comments, or others' comments here that we should look at existing literature, are inappropriate. But I would point out that, by definition, any of the literature that's over a couple of years old is not based on experience. There's a big difference between talking about theory and talking about experience. It's the experiences that we've really only been having now for a few years on which we're going to start building the literature. By all means, we should scan all the literature for concepts, but anything in the early literature must be suspect in the sense that it is not experientially based.

This is perhaps one of the really exciting times. The Chinese say it's a curse to live in exciting times, but I find it rather enjoyable--maybe it's making the best of a bad thing. But I rather enjoy a sense of turmoil, and I assume that most of you do, too, or you wouldn't be here.

One of the biggest failures in our discussions, I think, is not to make our sense of time explicit. It gets awfully tedious if we do it, but sometimes we had apparent conflicts when people said things, and it wasn't really a conflict. It's that this person is talking about the problems of today, and this person is talking about the problems five years from now.

I do think it's very important that we keep a kind of view forward. At the same time, I can understand that somebody like Ed Brownrigg has a greater concern with the problems of today than perhaps I would. I'm writing a dissertation, and he's serving thousands of users.

The situation is very volatile, and it's not just a question of our own evolution, but some people have pointed out the user also keeps changing. The user keeps changing, not only in getting a PC--I was in Wisconsin about a year ago and visited Waunakee High School, where they have a kindergarten-through-12 computer curriculum. The kindergarten one--evidently they start out (and I think it's a very bright idea), they start out by showing a lowercase letter on the screen, and the child has to hit the uppercase letter on the keyboard. It's computer literacy--fascinating. It also teaches them about upper- and lowercase letters.

We're going to be running up against a bunch of people coming in as freshmen very soon who are going to have very, very high expectations, and it's going to make a difference. It's not that we don't have high expectations at the moment, but once we solve those, there's going to be no rest for the weary.

I saw a number of trends in our discussion. One thing we're obviously seeing--again, breaking the either/or thing--we're seeing a simultaneous decomposition of systems and an increased linkage of systems. And I'm not

sure it was Ray or who it was that pointed it out, but these aren't necessarily in any way conflicting, as long as we understand, as long as we have a clear conceptual design of what's going on. A word that didn't come up that obviously applies is "modularity." As we move in to true network designs, as local area networks come along, it's our good old friend modularity that we were all taught about 15 and 20 years ago and learned about when we did our first coding.

We are facing a real rock-and-a-hard-place dichotomy between capability and capacity, because we are being pressed on both scores. We're being pressed to keep our costs down and, at the same time, the requirements for these systems are just exploding, and they're exploding in a variety of ways. It's my sense that, for the long term, we must optimize for flexibility. That also has been observed by a number of people here.

In terms of requirements, we have little arguments about them. We talk about online catalogs and we're really not talking about catalogs in any sense that we used to use that word. We're talking about public access to circulation files, public access to journal analytics. Is it pre-announcing, Ed, to say that MELVYL will be making some efforts in terms of providing journal analytics through its catalog in this next year, in cooperation with some other groups? We're going against in-process files, and then we're going against all of those special collections and uncataloged things.

I think in some large library systems, as little as 20% or 30% of the materials in the library system are really cataloged--they are all put off in side things. Our users are no l' ger willing to accept that. Karen Markey tells a delightful story of being berated by someone about 15 years old, as I recall. He said: "How DARE you put this thing up and not put the Readers' Guide in?" He really got angry and shouted.

We're also seeing, at the same time--and it's just exploding in every direction--we're seeing a need for increased personalization of systems. I'd like to suggest, as it has been suggested by a few others, that the personalization of systems will be according to the nature of the data.

Let me throw out an example. It seems to me that people who approach Shakespeare want to approach it maybe a little bit differently than somebody who's looking for a technical article. So we might well have, in our free enterprise society, somebody setting up a system that only examines the Shakespeare literature, works by and about Shakespeare, and permits all sorts of elaborate distinctions of editions and so forth. This would relieve the burden from our generalized systems for recording editions for the technical literature, where all we want is the latest one, where we are really not concerned about doing comparisons of texts.

Ed Brownrigg and others have pointed out that we're moving toward a situation where we have online texts that are constantly updated, and every time somebody prints a work out, it's going to be different from the last one. This is moving us back, in a sense, to the age of the scribes.

Carrying the Shakespeare analogy a little bit further, though--what the high school student wants to know about Shakespeare is a good bit different than what the Shakespearean scholar wants to know. So we'll probably end up with maybe two approaches to the Shakespeare literature. The kinds of things that will be given to the high school student, the kinds of help screens, especially, will be quite different than the sort of things we're going to give to the scholar.

Again, taking this tack of increased linkage and increased decomposition, I see us eventually moving toward a situation where we're providing a window on the world through our system. People are going to go out to whatever kind of approach to literature they want, on whatever machine it is on. It may well be that some of those approaches are sitting on somebody's PC somewhere. We'll link into some electronic cottage, and attack the database in there.

John Naisbitt in Megatrends has pointed out that the tendency, as we get a sort of monolithic communications access--and it may not be monolithic once you get inside it, but it looks that way to somebody approaching it--that the result on the other end is going to be that every person can be their own publisher. We're seeing that already in Blackwell North America; we're seeing the number of publishers expanding dramatically.

Emery Koltay of Bowker threw a number at me some time ago; I think he said they had given out 500 new publisher prefixes in the ISBN office in a single month, about two months ago. There is an explosion of publishers, and many of these publishers are doing single titles.

All of these kinds of explosions are going to mean that, as librarians, we're going to be more and more important in trying to help people negotiate this sea of publications. Word processing is also going to explode the number of publications that are out there, as well as world literacy, which is going up dramatically.

In addition, our users are being trained in a lot of ways--some of them like joy sticks, some of them like mice, some of them like roller balls. We're going to find ways to use these things. Clay Burrows was explaining to me that there are touch terminals that have a higher resolution than the current kinds of touch terminals that require a scored screen, really a sort of an artificial approach. These higher-resolution touch screens are expensive now, but they're going to get cheaper. We're going to reach a point where when you throw LISA-like icons up on the screen, somebody will say "that one!" to the voice input unit.

So we're going to have to look these new kinds of things for user input firmly in the face. It's not going to be the kind of thing where we really have choices. I mean, we'll have a choice--the choice will be to compete or not compete, be in the game or not be in the game.

Again, in this kind of environment, we must optimize for flexibility--that's really the most important thing we can do in terms of our institutions.

I would like to close with an observation Mike Gorman made, also at that last preconference. It's related to something that one of the other speakers here suggested, and I paraphrase--the demand curve may outrun our capacity curve. I think that's happening very quickly in a lot of places. Michael Gorman's line, and I think it's a good thought to end on, is that "There is a sleeping beast of demand out there, and the online catalog is going to waken it."

I think that if we come back five years from now, we're going to look at the problems we discussed and think how simple they were. Because the online catalog is not just automating the catalog. The online catalog is that last step that is going to cap all of the other efforts we've done in automation over these last couple of decades--and most of the people whom we are delivering online catalogs to don't understand that.

The demands that we're going to have are going to come in year after year and press our systems harder and harder, both in terms of quality and in terms of quantity.

QUESTIONS AND DISCUSSION

Question: Could you expand on your ideas about joy sticks and touch screens?

Aveney: The point really is that there is a whole variety of ways of expressing spatial movement, other than the scored touch screen, which is the only non-keyboard example we have had in our online catalogs, which gives you a very limited number of points.

When you can start showing the intersection of Boolean sets on a screen and touch the interception of the sets and say what you want--thanks to Clay Burrows for that idea--that's going to be a different kind of interface, and I think it's clear in my mind that before 10 years are up, we will be talking about non-keyboard colored graphic front ends.

Comment: Don't forget Millard's gear shift.

Aveney: I thought of the automatic choke when you said that, and I remember that when the automatic choke appeared, all of the Volkswagen owners were real concerned about how good that automatic choke was, and "it's sticking," and "it isn't as good as the old one we worked with," but somehow we all got used to that one, we got used to the controls.

Comment: Maybe the more quarters you put in, the faster response time gets. (laughter)

Question: You said that the online catalog is sort of a cap to the other automation we've done. I rather take exception to that. I think in the long run, the online catalog will be seen as a temporary aberration of technology at the time. We'll be all much happier when it's gone, simply because the online catalog supplies people with bibliographic citations, which is not really what they want, it's only intermediary. What they really want is data, and I don't think it's going to be that much longer before we can get rid of the idea of the online catalog, and have information management systems where data is given to people.

Aveney: That's already coming in, and I agree with that. I would still say--my sense of the cap is that it subsumes all of the others; it's the public face for all of the others. There is no question that the catalog is going to expand. One of the trends that Ed Brownrigg and I have talked about a lot is that there is going to be a move from the inventory approach, the online catalog as an inventory of a collection, to how you get access to things.

And at that point, the model is going to become a sales catalog, and when the model becomes a sales catalog, it's probably the seller who is going to supply the description. So we're going to see the end of cataloging departments in libraries, I think, within ten years. We've already seen them cut back dramatically.

Comment: I think that our systems are really much more temporary and ephemeral than we tend to think at any particular time. Any system that any of us has in operation or in planning is going to be obsolete and disappear in about 15 years, I would think, at the most. I think that we ought to keep that in mind.

Aveney: Anybody know an automated system that has run 15 years? We did a little thing in the Journal because I've had--my acquisition system has been running 13 years at Harvard, a batch system, it's a record-keeping system fundamentally. There are a few circulation systems out there, but I'm not aware of anything that's been running 15 years that's essentially the same system.

Comment: That's an interesting phenomenon in this area of computing that we're dealing with. I think systems typically last much longer in the library world than they do in, say, banking or elsewhere. I know some things I wrote in the late sixties that are still running. It's simply a matter of economics, probably.

Aveney: In our company, we amortize systems in five years, and that's it.

Comment: If I could just reiterate a comment that I made the other day--that I think we're going to, with the cost of these systems not going down, no matter what is happening--we are going to be forced into utilizing the software for longer and longer periods of time like they have in other industries. We just have to be flexible, and be able to take these systems of today and make them adaptable and usable for the systems of tomorrow, because--

Comment: And make them modular so we can...

Comment: We're not going to be able to afford to rewrite what we're doing right now. We will be building on it. We will be making new access and better systems. But the basic underlying functions we have today, we're going to need in some state...

Comment: I just want to pick up on what Millard said earlier. I was going to take issue also with the capstone idea. I speak as a non-librarian who has been associated with libraries and has been somewhat perplexed by them for a number of years. It seems to me that the library catalog is what it is today because of a large number of costs and technology factors like the size of three inches by five inches and so on.

As soon as the average user, the typical user, is going to be a kid sitting in his dormitory room, he's going to ask questions like, "Why is there only one subject classification code? Why isn't a table of contents in this record? Why aren't all the citations in this record?" Since most of this

stuff is available through other systems, it's not clear that the library catalog can be looked at as the capstone, but rather, just one more link in a very large chain.

Aveney: What I was trying to get across was a sense of spanning rather than a sense of topping off. It seems to me that either the library catalog is the access to all of the various information resources that someone is going to get to, or some other system will do that, and the library catalog will be one of the databases that this other system tacks in to.

Comment: I favor the latter mode, myself. Incidentally, 15-year-old systems: RECON.

Comment: Brian, I think implicit in your summary was a point that Charles made explicit with his sequence of transparencies, and that was a dramatic shift in the nature of the design process itself--and I think that we cannot underestimate the significance of that. It was back in the sixties, as I recall, one of the favored phrases when they were beating on the educationists at the time, that education was too important to be left to teachers.

There is the possibility that the notion of the retrieval system will become too important to be left solely to the designers, and in fact that has become the case, I think, as Charles has documented here. I think that we, to a certain extent, may resent some intrusions, but on the other hand, it is inevitable. The fact is that we leave out the key folks in that process at our own peril.

Whether they are the user or the reference librarian, the folks who are responsible for the training and education, whose contribution, I think, is underappreciated, or at least I suspect it has been to date. The fact is that it is a much larger, involves a much larger group of folks--

Aveney: My observation on that, in my own history in this business--when I started out, a library director, for reasons of prestige or assumed cost-benefit, would ask me to develop a system and sell it to the staff. I believe in the case of the online catalog that, by and large, it's the staff making demands. The library director is running to catch up and saying, "Can you deliver me something so I can meet this demand?"

So in a sense, we have gone from being producer-driven to being demand-driven--and again, the Gorman comment, that demand is going to go up dramatically. Our users are going to have tremendously rising expectations.

Comment: You said several times that often our access to contact with the user really is in the role of an intermediary and to carry another comment one step further, if you blow it, as a librarian, then people are likely to say that provision of the information society is too important to be left to librarians. There could be an awful lot of other kinds of intrusions at that

level, particularly in academic institutions. The electrical engineering department suddenly becoming instant experts on library automation--

Comment: They already think they are.

Comment: I want to argue with the point about, we're going to have to live with the software we write now, we can't afford to rewrite it. Like hell, we're going to be rewriting it over and over again in the next little while.

Aveney: It's going to be hard to rewrite the files as they get larger.

Comment: Well, you don't rewrite the files, you may change them and convert them, but not rewrite them.

Comment: They won't stay the same, though. We've got different tools now than we had 10 years ago in software development and database management. I think in 10 years it's going to change even more dramatically than it has in the last couple of years.

Comment: Well, there's no question that it's rare to see a 15-year-old system, which is another way of begging the question of where are we and where are we going? You can't talk about online catalogs without talking about the cataloging, and you can't talk about catalogs without talking about libraries, and you can't talk about libraries without talking about education.

I'll give you a fast forward, starting in the mid-18th century, when the process of education at a university involved professors who were the librarians, giving courses in the library and showing students where the material was. And for a long time, that process continued, but dwindled, where the professor was supposed to be the font of everything. When I was a graduate student at NYU, I got my education at the New York Public Library, and I couldn't have done it without the catalog there. And I think that was a silent, but very important, break in the tradition.

The locus of cataloging during our generation moved out of redundant cataloging all across the landscape into redundant catalog copying, with most of the source being Library of Congress and the rest being contributed by members.

The locus of cataloging is still moving, and I think this is what Brian is telling us to keep our eye on. It is moving, and I think it's moving more and more to the publishers who are going to use it as an advertising medium. This is because, as more than one of us has observed in here, the users don't want citations, they want the data. How you advertise the data is going to determine its availability, its quantity and, therefore, the profit in disseminating it. That's the business we're in.

Aveney: It's the same phenomenon as robots coming in to automobile plants, it seems to me. We're getting identification of those tasks that require human beings and those that are highly repetitive. We know there are professional and clerical sides to librarianship. I would suggest that we're going to see a lot fewer librarians around. We're going to see the clerks essentially automated out. What we're going to see then is increasing professionalization of those librarians who are around, really serving people, talking to the public, helping them negotiate things. What we may well see is librarians survive, and libraries not. And that's one possible scenario.

Comment: Ed's remark leads me to let this group know that the Council has just recently provided $50,000 to the Association of American Publishers, to help them create a standard for coding manuscripts to be submitted in electronic form for publication. One of our motives for becoming involved in an environment that clearly produces a lot of profit, and might well have been expected to fund that on their own, is to provide an entry point, if you will, for the library world.

Some of those codes we want to be codes that identify author-generated descriptors of what they produce, instead of third-party-generated descriptors. Whether or not this effort results in a standard that is used is sort of beside the point. The real point is that the effort is being made. Something along these lines is going to happen. If the library world isn't involved at the point of specification, in the establishment of the standard, we could indeed be left holding our MARC standard with nothing to put in it to shove around the countryside.

So we're making sure that at least the library world has some input into what those codes will be, the resulting output data that will be available. And it is indeed quite possible. A lot of the data that we now believe or describe as cataloging data--it may not fit any of our present formats--will be generated at the publisher level, downloaded into the library system, and used quite differently than the records we now are familiar with.

Aveney: I've got to just put a gloss on that to say that there are identical trends going on in publishing, in librarianship, indeed in general; what we're finding--decentralized computing power, decentralized information--will eventually lead to decentralized political and economic power.

And just as it seems to me that we're going to see increasing emphasis on the individual professional skills of librarians, and less reliance on an organization that has a body of capital resources, on procedures, and so forth--we're going to see the same thing happening in publishing. Again, we're seeing a lot of individual publishers. We're going to move, you know, if Toffler and Naisbitt and a lot of the other people are correct, toward a society in which we have much more decentralized power, not necessarily the way it's used in foreign policy and so forth, but in terms of the ability of individuals to do things. This is going to mean that sorting our way through this world is going to be more difficult because they won't clump up into these large industrial model institutions.

Comment: I had two comments I wanted to mention. The first one is that, what is so interesting about the last 12 months for me is that the decision has been to redirect the investment or to balance the investment we have made in processing support, with an investment made in service support. Which is to say, in fewer words, that the money we're spending now on technology, on this patron access catalog, is visible to the public. To this point in time, that has not been strictly the case in an institution like the one I work at.

I'm not sure the public ever noticed that the Xerox 9700 was printing their catalog cards. They're certainly going to notice the next million dollars we spend at Columbia on bibliographic control--because it's not really being spent on bibliographic control.

The second thing I wanted to say is to echo, from my perspective, this idea of the passing of the cataloger or the passing of the reference librarian, but to come at it from a slightly different angle.

At Columbia, we spend about $4,000 a year in capital support per professional staff person, and I think that this is the same sort of economic analysis that is frequently made in production enterprises. What is your labor/capital ratio?

We are very, very low right now--throughout the entire industry, with respect to office, let alone intellectual community, scholarly publishing endeavors.

And whether we're going to have more money spent per professional, more capitalization per professional by virtue of a decrease in professionals, or an increase in spending--I don't know--it could go both ways. But what is happening now is that we're operating under "more is better." More capital per professional worker is better; whether that professional worker is called a librarian, I don't know. Right now I'm just working on the problem of breaking down the distinction between cataloger, bibliographic searcher, acquisition specialist, bibliographer, reference librarian, etc. At the same LITA preconference, Brian Nielsen gave us the most information on this, and I recommend that his paper be read in light of this.

But the second part of it is the capital per professional worker--and I think that is a very important cost-benefit perspective on what's happening right now.

Comment: I'd like to comment on the last statement. I read an article a few years ago about an economist who was pointing out the general esteem with which one's profession is held in the general public. Interestingly enough, he ran a coordination between the esteem level and how many dollars were spent for technology and so on in support of their profession. Medicine, of course, was way up there. And it's that sort of thing I think that we're talking about.

Comment: It seems to me, too, that there is a shift in the sense of automation. In tech services, it's largely to do the same functions faster, or to do more of them with the same manpower. Whereas the public service side of it seems to me to be more of an intellectual thing. I keep wondering how much, as the public interface develops, and I see the keyword in Boolean being maybe a first step, and then artificial intelligence--the possibility seems to go even farther toward eliminating some of the background work and authority control in maintaining the database. That is now the most significant cost element. How much of that can be eliminated by simply having the front end sort some of it out?

Comment: In terms of the cost of the investment per individual, I think it's a question of productivity. I read somewhere that agricultural workers are supported by $40,000 worth of capitalization, and industrial workers--depending on the industry they're in--are supported by $25,000 to $30,000 worth of capitalization per worker. Intellectual workers are supported by less than $2,000--which is a table and a typewriter.

Comment: I think one of the things that happens with our institutions--both the public library system and the academic--is that, in effect, the library has called on a certain percentage of the resources available, and that's money, in exchange for services and supplies. The online catalog is putting us in a situation in which the library is sort of saying to the community, "Give us more resources because we want to supply better services." That's not a decision that the library vendors are ever going to make. It's not even one that just libraries are going to make.

You know, people walk into the library who never consciously realized that the library is 15% of the university's budget, or whatever it is, and that they get service for that--because it's all free for the user. Suddenly they're going to start seeing it. Now, if the service is worth it, they're going to pay more. I think what we have to be very careful about is that the libraries don't commit themselves to paying more before they convince their clientele that it's worth doing that.

Comment: Incidentally, and I don't want to blow your one figure--but 15% is a little wide. We get 5.3, and Harvard gets 9.

Comment: I'm not too comfortable with the assumption that everything that librarians do in libraries could be eliminated by good electronic systems. I think we should pay a little more attention to those sort of tasks that are best handled by people. For instance, when we talk about help screens, nobody ever suggested that, at some point, the face of a librarian should flicker on the screen and should ask the person what they want.

Now that's not as crazy as it sounds. I think a lot of the times some professional intuition would be the best thing to put into the online catalog, and as far as I know, it's always 100% of the goal to eliminate librarians.

Aveney: Put broadband lines in and switch to live video, in which the librarian comes up on your CRT and says, "You seem to be having a problem here--can I do anything to help you?"

Comment: In fact, some new parts in our system specifically suggest to the user that they go talk to the librarian.

Comment: That's what we did, and all the librarians are complaining.

Comment: This has been very interesting because I learned a lot--if nothing else, I've learned that if I hang too long on a single branch, I'll probably fall on my own triple-edged sword. (laughter)

The thing that really is mind boggling is just the thought of what may happen in 5 and 10 and 15 and 20 years. But for many of us, I think we are faced with many more immediate problems than what's going to be happening 15 years from now.

I'm not saying that we don't have to look at those; I think we do--because we do have to anticipate and we do have to be ready to change. But I still struggle with a basic thing--with all of the change that is going to be happening and is happening, I'm still not certain that, in even half of the situations, we're going to be able to give everything to everybody. We're not going to be able to give them all they want. We are not going to be able to have such tremendous variety and also such tremendous consistency that they're going to be mobile and they're going to go anywhere and they're going to feel at home.

Now I'm not really speaking against standards, but it's going to be one heck of a long time before we're ever going to solve that problem of being flexible, modular and consistent and simple and--I just don't see it happening it that way.

Aveney: We're not really writing specs right now when we talk about these kinds of things. It seems to me what we're doing is giving a weather warning, and weather warnings are correct some percentage of the time. We're saying, you know, maybe you still want to go out to the ocean, but you may want to take an extra slicker or so on. The point of these kinds of discussions, it seems to me, is so that we do think in terms of modularity and we do think in terms of separating off those modules that we think there is going to be the greatest change in.

Because, yes indeed, we are going to have to rewrite, but we can't constantly keep going and reconstructing. We can't keep rebuilding the architecture of systems. And so again, that's my sense, that optimizing flexibility has to be the approach that we take in terms of the long-term economic interest. And people have said that in different ways.

APPENDICES

APPENDIX A

AGENDA

ONLINE CATALOG DESIGN ISSUES
A Series of Discussions

Holiday Inn - Inner Harbor
Baltimore, Maryland
September 21 - 23, 1983

WEDNESDAY, SEPTEMBER 21, 1983

6:00 PM	Introductory Libations	Center Ballroom
6:45	Supper	
7:30	Introductions	Center Ballroom
8:00	Discussion Topic #1 John Schroeder - File Structures and Computing Resource Management	

THURSDAY, SEPTEMBER 22, 1983

7:45 AM	Continental Breakfast	West Ballroom
8:00	Discussion Topic #2 James Corey - Search Facility Options; Boolean (Implicit/Explicit), Keyword, etc.	West Ballroom
9:30	B R E A K	
10:00	Discussion Topic #3 Charles Hildreth - User Feedback Mechanisms	
11:30	LUNCH ON YOUR OWN	
1:30 PM	Discussion Topic #4 Joe Matthews - Screen Layouts and Displays	West Ballroom
3:00	B R E A K	
3:30	Discussion Topic #5 Michael Monahan - Command Languages and Codes	

5:00	Rest and Relaxation	

THURSDAY, SEPTEMBER 22, 1983 (continued)

6:00	Cocktails	Center Ballroom
6:45	Supper	
7:30	Discussion Topic #6 Clay Burroughs - Online User Prompts and Aids	West Ballroom

FRIDAY, SEPTEMBER 23, 1983

7:45 AM	Continental Breakfast	West Ballroom
8:00	Discussion Topic #7 Ray DeBuse - Links with Other Systems, Internal and External	West Ballroom
9:30	B R E A K	
10:00	Discussion Topic #8 Edwin Brownrigg - Telecommunications and Online Catalogs	
11:45	LUNCH AT ZIGGIES	Ziggies
1:00 PM	Summary of Discussions Brian Aveney - Proceedings Editor	West Ballroom
1:30	Concluding Discussion Comments, Suggestions, Recommendations	
3:00	HAVE A SAFE TRIP HOME	

APPENDIX B

LIST OF PARTICIPANTS

ONLINE CATALOG DESIGN ISSUES
A Series of Discussions

Holiday Inn - Inner Harbor
Baltimore, Maryland
September 21 - 23, 1983

1. J A M E S A A G A A R D
312-492-7641
Northwestern University Library
1935 Sheridan Road
Evanston, Illinois 60201

2. B R I A N A V E N E Y
503-684-1140
Director for Research and Development
Blackwell North America
6024 S.W. Jean Road, Building G
Lake Oswego, Oregon 97034

3. E D W I N B R O W N R I G G
415-642-9485
University of California Systemwide Administration
Division of Library Automation
186 University Hall
Berkeley, California 94720

4. C L A Y B U R R O W S
206-786-1111
Biblio-Techniques Inc.
828 East 7th Avenue
Olympia, Washington 98501

5. D A L E C A R R I S O N
507-389-6201
Dean of the Libraries
Mankato State University
Mankato, Minnesota 56001

6. V I N O D C H A C H R A
703-961-6122
Director for the Center for Library Automation
Vice President for Computing and Information Systems
Virginia Tech
201 Burruss Hall
Blacksburg, Virginia 24061

7. P A U L A C H E S S I N
617-965-6310
CLSI Inc.
81 Norwood Avenue
Newtonville, Massachusetts 02160

8. J A M E S C O R E Y
314-882-7233
Director, Office of Library Systems
University of Missouri
523 Clark Hall
Columbia, Missouri 65211

9. R A Y D E B U S E
206-459-6522
Director of Planning and Development
Washington Library Network
AJ-11
Olympia, Washington 98504

1 0. R I C H A R D S. D I C K
301-983-8900
Executive Vice President
Avatar Systems, Inc.
11325 Seven Locks Road, Suite 205
Potomac, Maryland 20854

1 1. T O M D O S Z K O C S
202-496-6531
National Library of Medicine
8600 Rockville Pike
Bethesda, Maryland 20814

1 2. E M I L Y F A Y E N
603-646-2235
Dartmouth College
Baker Library
Hanover, New Hampshire 03755

13. DOUGLAS FERGUSON
415-497-9724
Systems - Green Library
Stanford University Libraries
Stanford, California 94305

14. ERIC FERRIN
814-865-1818
E 8 Pattee Library
Pennsylvania State University
University Park, Pennsylvania 16802

15. MARK FOSTER
608-263-9197
University of Wisconsin
1120 West Johnson Street
Madison, Wisconsin 53706

16. JEFF GRIFFITH
202-287-6447
Congressional Research Service
Library of Congress
Washington, D.C. 20540

17. CHARLES HILDRETH
614-764-6172
OCLC, Inc.
6565 Frantz Road
Dublin, Ohio 43017

18. LYNDON HOLMES
617-965-6310
CLSI Inc.
81 Norwood Avenue
Newtonville, Massachusetts 02160

19. MILLARD JOHNSON
314-454-3711
Washington University
School of Medicine Library
4580 Scott Avenue
St. Louis, Missouri 63110

20. JOE MATTHEWS
916-272-8743
J. Matthews and Associates, Inc.
213 Hill Street
Grass Valley, California 95945

21. A. STRATTON McALLISTER
0711-7839-1, X4883
IBM - Systems Service
Vaihinger Strasse 151
D-7000 Stuttgart, 80
WEST GERMANY

22. CARYL McALLISTER
0711-7839-1, X4883
IBM - Systems Service
Vaihinger Strasse 151
D-7000 Stuttgart, 80
WEST GERMANY

23. CHARLES MEADOW
415-858-3783
DIALOG Information Services, Inc.
3460 Hillview Avenue
Palo Alto, California 94304

24. JAMES MICHAEL
800-325-0888
Data Research Associates
9270 Olive Boulevard
St. Louis, Missouri 63132

25. MICHAEL MONAHAN
416-475-0525
Geac Computers International
350 Steelcase Rd. West
Markham, Ontario
CANADA L3R 1B3

26. TED MORRIS
312-962-8763
University of Chicago Libraries
Joseph Regenstein Library, Room 210
1100 East 57th Street
Chicago, Illinois 60637

27. PAUL PETERS
212-280-4744
Assistant University Librarian for Systems
Columbia University Libraries
215M Butler Library
New York, New York 10027

28. NOLAN POPE
904-392-0796
505 Library West
University of Florida
Gainesville, Florida 32611

29. LARRY PORTER
519-824-4120
University of Guelph Library
Guelph, Ontario
CANADA N1G 2W1

30. WILLIAM POTTER
217-333-0318
University of Illinois
246 A Library
1408 West Gregory
Champaign-Urbana, Illinois 61801

31. JAMES SCANLON
404-542-2716
Manager of Library Automation
Office of Computing and Information Services
University of Georgia
Athens, Georgia 30602

32. JOHN SCHROEDER
415-328-0920
Director of Research
Research Libraries Group
Jordan Quadrangle
Stanford, California 94305

33. WARD SHAW
303-571-2270
Colorado Alliance of Research Libraries
c/o Denver Public Library
3840 York Avenue
Denver, Colorado 80205

CLR STAFF

C. Lee Jones, Program Officer

Keith Russell, Program Associate