Volume 10,
Number 1

ISSN 0730-9295

Information Technology and Libraries

March 1991

CONTENTS

If your online catalog contains misspellings, typographical errors, and old forms of headings that confuse and frustrate catalog users, contact Blackwell North America, the leader in automated authority control for over 15 years. Blackwell's authority control service, based on Library of Congress name and subject authority files, standardizes headings for improved online access. Write or call today for more information on Blackwell's Authority Control Service.

Technical Services Division
BLACKWELL
NORTH AMERICA, INC.

Technical Services Division, Blackwell North America, Inc., 6024 S.W. Jean Road, Building G, Lake Oswego, Oregon 97035
Phone Toll Free: 1-800-547-6426 or 503-684-1140 (in Oregon) FAX: 1-503-639-2481 Canadian Toll Free: 1-800-626-1807

ISSN 0730-9295

Information Technology and Libraries

Volume 10, Number 1: March 1991

CONTENTS

ITLBDC 10(1) 1–84 (1991); Inf Technol Libr

Information Technology and Libraries

Information Technology and Libraries is the official publication of the Library and Information Technology Association, a division of the American Library Association, 50 E. Huron St., Chicago, IL 60611; *Executive Director*: Linda J. Knutson. The journal is issued quarterly in March, June, September, and December.

Information Technology and Libraries publishes material related to all aspects of library and information technology. Some specific topics of interest are: Automated Bibliographic Control, AV Techniques, Communications Technology, Cable Systems, Computerized Information Processing, Data Management, Facsimile Applications, File Organization, Legal and Regulatory Matters, Library Networks, Storage and Retrieval Systems, Systems Analysis, and Video Technologies. *ITAL* welcomes unsolicited manuscripts. Submissions should follow the guidelines stated under "Instructions to Authors" on page 84 of the March 1991 issue. Manuscripts of articles, communications, and news items should be addressed to the editor: Thomas W. Leonhardt, *Information Technology and Libraries*, University of the Pacific, Stockton, CA 95211. Copies of books for review should be addressed to: Susan B. Harrison, *ITAL Book Reviews*, The New York Public Library, 455 Fifth Ave., New York, NY 10016. Copies of software for review should be addressed to: George S. Machovec, *ITAL Software Reviews*, Arizona State University, Tempe, AZ 85287. Advertising arrangements should be made with Kurt Murphy, Hayden Library, Arizona State University, Tempe, AZ 85287.

Information Technology and Libraries is a perquisite of membership in the Library and Information Technology Association. Subscription price, $20, is included in membership dues. Nonmembers may subscribe for $40 per year in the U.S., Canada, Mexico, Spain; $45 in other foreign countries. Single copies, $11. Second-class postage paid at Chicago, Illinois, and at additional mailing offices. *Postmaster*: Send address changes to *Information Technology and Libraries*, 50 E. Huron St., Chicago, IL 60611.

Circulation and Production: ALA Publishing Services (Eileen Mahoney; Dianne Rooney; Amy Brown, Bruce Frausto, Josephine Gibson-Porter, Dan Lewis, and Donavan Vicha), American Library Association, 50 E. Huron St., Chicago, IL 60611.

Publication of material in *Information Technology and Libraries* does not constitute official endorsement by the Library and Information Technology Association or the American Library Association.

Abstracted in *Computer & Information Systems*, *Computing Reviews*, *Information Science Abstracts*, *Library & Information Science Abstracts*, *Referativnyi Zhurnal*, *Nauchnaya i Tekhnicheskaya Informatsiya*, *Otdyelnyi Vypusk*, and *Science Abstracts Publications*. Indexed in *CompuMath Citation Index*, *Computer Contents*, *Computer Literature Index*, *Current Contents/Health Services Administration*, *Current Contents/Social Behavioral Sciences*, *Current Index to Journals in Education*, *Education*, *Library Literature*, *Magazine Index*, *New Search*, and *Social Sciences Citation Index*. Microfilm copies available to subscribers from University Microfilms, Ann Arbor, Michigan.

The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences–Permanence of Paper for Printed Library Materials, ANSI 239.48-1984. ♾

User Problems with Access to Fictional Characters and Personal Names in Online Public Access Catalogs

Martha M. Yee and Raymond Soto

Reference librarians in libraries with online public access catalogs were polled to determine what indexes they thought most patrons would use to look for fictional characters. Most thought users would choose a name index over a subject index, but that they would choose a subject index over an author index. In addition, reference librarians were asked their own preferences. They said that the best kind of index to give users for such searches would be a general index in which no particular kind of search need be specified. Based on the descriptive statistics from this research, current system design practices for online public access catalogs and current tagging practice in the MARC format are evaluated, and recommendations are made for modification of those practices.

An examination of the forty-eight online public access catalogs described in Matthews' *Public Access to Online Catalogs*, 2nd edition[1] reveals that all but four require users to choose a particular index prior to conducting a search. Most of these catalogs offer the user a choice between an "author" index and a "subject" index. Only four offer the user a choice between a "name" index and a "subject" index. Four other systems do not require the user to choose a particular index; instead, all searchable fields are searched for a match with input search terms.

The choice between author and subject indexes or between name and subject indexes can be a difficult one for users who are searching for fictional characters. Take, for example, a user who is searching for any available Sherlock Holmes film or book. When the user is presented with the choice of a name index or a subject index, he might choose the name index. Currently, fictional characters are tagged in the MARC format[2] as topical subject headings (650 field). Since there is no way for a machine indexing program to distinguish fictional character names from other topical subject headings, there is no way for systems to include fictional character names in a name index. Thus, they must be indexed in the same way that other topical subject headings are indexed. When presented with the choice between an author index or a subject index, the user might be stymied, since Sherlock Holmes is clearly not an author and might not seem to be the "subject" of a fictional work either.

Martha M. Yee is Cataloging Supervisor at the UCLA Film and Television Archive. Raymond Soto is Public Services Librarian at the Theater Arts Library, University of California, Los Angeles.

The research reported in this article was supported by a grant from the Librarians' Association of the University of California, for which the authors would like to express their gratitude. The authors would also like to acknowledge the help they received from Allyson Carlisle, Rita Engelhardt, Sara Shatford Layne, and the public service librarians of UCLA and Thousand Oaks Public Library (Thousand Oaks, California) for their help with questionnaire design, statistical review, and editing of the manuscript.

To the authors' knowledge, no research has yet been done on potential difficulties that users might have when choosing among the indexes offered in various online public access catalogs. Recent articles by Ross, Sanders, Klugman, and Pilachowski and Everett have discussed possible difficulties, but these articles have not reported any research results.[3] None of these articles discusses possible problems presented by fictional characters.

By means of tagging, the MARC format identifies a number of potentially indexable categories (see table 1). While each of these categories is available for separate indexing, most systems offer only name, title, or subject indexes. Each index combines several MARC-tagged categories. The development of these categories, both of the MARC format and of existing online systems, was intended to fulfil the system needs and the needs of catalogers. To the degree that MARC and the systems that use it have attempted to address the needs of end-users, intuition rather than research has been employed in decision making about system design.

Recently, the Subject Analysis Committee (SAC) of the Association for Library Collections & Technical Services (ALCTS), which is part of the American Library Association (ALA), formed a subcommittee to produce national standard guidelines for the provision of subject access to individual works of fiction. At the ALA Midwinter Meeting in 1990, the standards were approved and subsequently published by ALA.[4]

Table 1. Indexable Fields Currently Available in MARC Format.

600	Subject added entry—personal name
610	Subject added entry—corporate name
611	Subject added entry—meeting name
630	Subject added entry—uniform title
650	Subject added entry—topical term
651	Subject added entry—geographic name
655	Index term—genre/form
656	Index term—occupation
657	Index term—function
700	Added entry—personal name
710	Added entry—corporate name
711	Added entry—meeting name
730	Added entry—uniform title
740	Added entry—variant title
752	Added entry—hierarchical place name
753	Technical details access to computer files
754	Added entry—taxonomic identification
755	Added entry—physical characteristics

Although the Library of Congress does not currently provide access by means of fictional characters to individual works of fiction, the subcommittee's guidelines recommend providing such access. The subcommittee also recommended that the Library of Congress attempt to provide fictional character access to current fiction, although it is not yet known whether the Library of Congress will be able to follow this recommendation. The research described in this paper was designed to give more detailed information about user needs in this area in support of a national effort to improve subject access to fictional materials.

It can be argued that fictional character access through indexes to online public access catalogs is part of a larger problem. Fictional characters are not the only entities that do not fall neatly into one of two or three broad categories that are currently offered for searching. There are other similar problems that may need investigation. Systems that allow users to search author/name and subject authority files are undoubtedly indexing uniform titles in the author/name authority file. It is doubtful that users know to look for them there.

Proper names as subjects can be given any of the following tags in the MARC format: 600, 610, 611, 650, 651. Even catalogers need help deciding which tag to use for a given name. Thus, catalogers consult a Library of Congress rule interpretation which is fondly known as the "dividing the world document."[5] According to this rule interpretation, the Empire State Building is tagged as a corporate name (610), and Ellis Island Immigration Station is tagged as a topical subject heading (650). This means that the former is indexed in the author/name index and the latter in the subject index. Again, it is doubtful that users know these differences. In the MARC format, geographic names are sometimes tagged as corporate names (when they correspond to governmental jurisdictions) and thus are indexed in most systems in the name/author index. Sometimes they are tagged as geographic names (when they are strictly geographical) and thus are indexed in the subject index in many systems.[6] Given the somewhat arbitrary scatter that librarians create by this kind of tagging, users may have

trouble deciding how to search for proper names in general. Unless a system has a general index, there frequently is more than one place to look for proper names given the current structure of the MARC format.

Finally, users may have problems searching for names as subjects, since some systems index them with names/authors, and other systems index them with subjects. Users might be expected to look for them in either place.

Since the authors of this paper were studying access to real people in fiction, they decided to include one question on nonfiction about real people just to see how the answers differed. They hoped this one question might provide some clues as to problems users may have with access to names as subjects.

DESIGN AND METHODOLOGY

The authors polled reference librarians rather than library patrons for several reasons. First, reference librarians have constant contact with users, training them in searching strategies, conducting searches on their behalf, and assisting users with difficult or failed searches. Therefore, we can assume that reference librarians have some insight into user needs and behavior. Secondly, reference librarians belong to professional organizations, which makes sampling easier. Naturally, the authors would have preferred to study user behavior directly, but if they had attempted to question library patrons, their sampling would likely have been confined to their own institution. Thus, the findings would not have been representative of the many kinds and sizes of libraries and types of patrons. Thirdly, it can be difficult to develop questions using terms that naive users can understand or interpret in the same way that the questionnaire designers or other users interpret them. Reference librarians are more familiar with the problems being addressed and are more likely to understand questions on a questionnaire. Therefore, the questionnaire can give more reliable and accurate results.

A questionnaire was designed and tested. It contained five questions concerning where users would be most likely to look for fictional and nonfictional characters in fictional and nonfictional works.

The authors sent the questionnaires to the heads of reference of 556 U.S and Canadian public and academic libraries that were likely to collect fictional materials and that are institutional members of the Library Information Technology Association (LITA) of the ALA. Since membership includes a subscription to *Information Technology and Libraries*, a major source of news and information about online public access catalogs, this would seem to be a viable assumption. The authors excluded business, corporate, law, or science libraries from this study on the assumption that such libraries do not serve users of fictional materials. Each head of reference was asked to give the questionnaire to a librarian who was likely to encounter users of fictional materials. The authors had a response rate of 85% of questionnaires returned, and 79% of questionnaires filled out and returned. Thirty-three percent of nonrespondents were in public libraries, and 62% were in academic libraries. These proportions are close to the proportions of respondents in both types of libraries; thus, there is no evidence of bias in the results due to a response failure of one type of institution. This impressive response rate is significant in its own right, in that it seems to indicate a strong concern on the part of reference librarians about the design of online public access catalogs and a willingness to be consulted. Table 2 provides data about respondents, their libraries, and catalogs.

FINDINGS

When users are looking for fictional characters, most reference librarians think they would choose name searches over subject searches, and subject searches over author searches (see tables 3 and 4). This means that the current tagging of fictional characters as subjects probably works better (i.e., is less confusing for users and elicits the correct choice on their part) in the majority of systems that call the index an "author index" than it does in the minority of systems that call the index a "name index." One might speculate, though, that in either type of system, the correct choice would not be self-evident to most users. Despite the fact that reference librarians thought the choice was less confusing when the index was called the author index, many of them chose to recommend the use of a name index (see table 5). Most reference librari-

Table 2. Characteristics of Respondents, Their Libraries, and Catalogs

Characteristic	%
No. of years as a reference librarian (N = 429)	
1–10	47
11–20	43
21 or more	10
Frequency with which users seeking fictional materials are encountered (N = 432)	
Once or more per day	30
1–4 times a week	27
Less than once a week	42
Proportion who think their patrons would find it useful to be able to search under fictional character names (N = 434)	
Not at all useful	5
Somewhat useful	42
Useful	27
Very useful	27
Types of library (N = 437)	
Academic	67
Public	30
Other	3
Size of libraries (N = 412)	
100,000 or fewer volumes	16
100,000 to 1 million volumes	62
More than 1 million volumes	22
Types of catalog (N = 438)	
Card	73
Microform	19
Online	55
Other (e.g., CD-ROM—3%)	9
Searches available on online catalog (N = 204)	
Author	93
Name	10
Subject	95
General search that does not require specifying search type	26
Search used for names as subjects (N = 202)	
Catalogs index as a subject in the subject index (e.g., works about Abraham Lincoln)	93
Index a name as a subject in the author/name index	7

ans think that the subject index is not a very good place for fictional characters and that the best solution would be a general index that does not require the specification of any index in the search (see table 5).

A number of people wrote passionate statements in the margins of their questionnaires about how difficult it is for users to deal with the categories created. For example, "people...do not know what author or subject mean, much less the differences between fiction and nonfiction or biography and autobiography." Several commented that a fictional character index would be more confusing than helpful to users. Reference librarians give different answers when the character is real rather than fictional (e.g., Lincoln). A name index was preferred for Lincoln in fictional works (between 27% and 35% at a 95% confidence interval), and a "works about" index was preferred for nonfiction works about Lincoln (between 27% and 35% at a 95% confidence interval). Chi square tests revealed no significant association between the answers reported above and characteristics of the respondent, type of library, or type of catalog available.

Some caveats about the findings: those who think users will look for Sherlock Holmes under author (see table 4) may have been slightly inflated by those who misunderstood and thought the question asked how many would look under Sir Arthur Conan Doyle. One respondent felt that librarians would be more likely to check the subject index than users would.

Table 3. Where Reference Librarians Think Users Will Look When Choice Is between Subject Index and Name Index

	Sherlock Holmes as character (book) (%) (N = 433)	Bugs as character (nonbook) (%) (N = 434)	About Bugs (book) (%) (N = 434)	Lincoln in fiction (book) (%) (N = 435)	Lincoln in nonfiction (book) (%) (N = 434)
Users will look under subject	12	18	54	19	37
Users will look under name	88	88	46	81	63

Table 4. Where Reference Librarians Think Users Will Look When Choice Is between Subject Index and Author Index

	Sherlock Holmes as character (book) (%) (N = 431)	Bugs as character (nonbook) (%) (N = 434)	About Bugs (book) (%) (N = 438)	Lincoln in fiction (book) (%) ()N = 437)	Lincoln in nonfiction (book) (%) (N = 435)
Users will look under subject	83	96	97	92	93
Users will look under author	17	4	3	8	7

Table 5. Type of Index Recommended by Reference Librarians

	Sherlock Holmes as character (book) (%) (N = 433)	Bugs as character (nonbook) (%) (N = 431)	About Bugs (book) (%) (N = 432)	Lincoln in fiction (book) (%) (N = 432)	Lincoln in nonfiction (book) (%) (N = 436)
Author	2	0	1	1	1
Name	28	22	18	31	24
General	37	39	32	26	21
Works about	7	6	21	16	31
Fictional character	23	27	15	17	0
Subject	3	5	13	8	21
Other	1	1	1	2	0

It was apparent from some of the respondents' comments that not all online catalogs are public access catalogs. The authors should have specified for respondents to include CD-ROM catalogs as well; many of the questions about online catalog searches could have been asked equally well about CD-ROM searching software. The general searches available may differ quite a bit from each other. Many checked the general search box when they had what they referred to as a "keyword" search. It is not clear, however, what might be indexed in a keyword search. Some keyword searches may search only titles

but not subject headings (where fictional characters would be found) or names. Also, some keyword searches may require specification of indexes (e.g., keyword subject search vs. keyword name search).

DISCUSSION

The authors' findings indicate that users are probably having difficulty choosing the correct index or type of search in systems that require such a choice. Currently, catalogers divide indexed terms in cataloging records into three broad categories generally referred to as titles, subjects, and authors. Online catalog designers create indexes based on these categories, usually requiring users to specify an index in a search. Unfortunately, there are types of headings that do not fall neatly into one of these broad groups. Fictitious characters are just one example of such headings. It is likely that a user looking for one of these types of entities (e.g., a fictitious character) will have difficulty deciding which type of index to pick.

It could be argued that the ideal solution to the problem of fictional characters (and other entities with proper names that fall into online catalog indexes in ways that are unpredictable to the hapless user) is to create a general index. This probably could and should be done in smaller catalogs. If, however, systems cannot afford the general index, which would be extraordinarily large, librarians must ask whether they need different categories for indexed terms, or whether terms currently indexed need to be put into different categories. In other words, they must ask whether change in the MARC format is required for a solution.

One solution that can always be considered is that of double indexing. For example, fictional characters could be indexed in both the name/author and the subject index. For this to be done, categories of terms that were candidates for double indexing would have to be separately tagged or labelled in some way. Otherwise, computer programs would have no way of distinguishing fictional character names, or other candidates for double indexing, from other purely topical subject headings that should not be double indexed.

Currently, library practice is to divide access point fields into three categories based on the MARC format: the 6XX fields, the 1XX/7XX fields, and the 8XX fields. The 8XX fields, designed for series, will not be discussed here. The 1XX, or main entry fields, while related to the 7XX fields, will not be be discussed either. The 6XX fields are titled "Subject added entry," and the 7XX fields are titled "Added entry." No statements are made explicitly about the principles behind the division into 6XX and 7XX fields, other than the general statement in the *USMARC Formats: Underlying Principles*[7] that the first character of the tag "identifies the function of the data within a traditional catalog record (such as main entry, added entry, subject entry)." One could roughly differentiate them as follows: 7XX fields are generally formulated based on the descriptive part of the cataloging record, including the physical description, while 6XX fields are based on a judgement of the cataloger as to what the work is about. As more and more indexable fields are added to the MARC format to accommodate various kinds of special materials, the distinction between 6XX and 7XX becomes more difficult to describe logically. For example, it can be hard to draw the distinction between the 655 and the 755 field. However, the original distinction between works by and works about is a valuable one to maintain, as it allows users to include or exclude one or the other category at will. Thus, a user who knows that he wants works by George Eliot, but not works about her, can narrow his search appropriately.

Under the 700 field, "Added entry—personal name," the following description is given: "This field contains personal names associated with the work which are not used as the main entry. This category includes (1) names of actual persons, capable of authorship, but not used as a main entry, e.g., editors, translators, etc., and (2) titles of works entered under personal author that are added entries." Thus, for deciding if a personal name belongs in the 700 field, the only real criterion the MARC format offers is whether or not the name is "capable of authorship."

It could be argued that one of the fictional characters in the study, Bugs Bunny, is a performing fictional character. *AACR2* recognizes most performers as authors. Thus Humphrey Bogart's name is placed in the 700 field for one of his movies. However, as noted above, Bugs Bunny goes into a topical subject heading 650 field. It seems conceivable that

some users, knowing that Bogart was in the name/author index, would expect to find Bugs Bunny there, too. Performing animals are also placed in 650 fields.[8] Thus, one unwritten rule seems to be that only human beings go into 700 fields; no animals are allowed.

The fact that performers are placed in 700 fields, and therefore placed in indexes called "author" indexes in most online public access catalogs, may be another source of confusion for the users. Would a user looking for Humphrey Bogart films realize that the author index should be chosen for his or her search? The authors suspect that most users would consider the term "author" to refer only to a person who writes a book and not necessarily to the artists, photographers, performers, etc., who are placed in 700 fields in current cataloging practice.

The word "name" may cause users problems as well. The authors work with an online public access catalog that calls its three indexes "name," "title," and "subject." One of the authors found users searching for journals in the name index because they were using the name of the journal in their search. Since uniform titles are indexed in the name index in the system, the user could have run across titles that would reinforce this practice and have failed to realize that he or she was missing journals that do not have uniform titles. The authors have already alluded above to the fact that many proper names will be in the subject index, not the name index.

No one would argue that fictional characters other than performing ones are capable of authorship and that is probably the basis on which fictitious characters are excluded from the 700 field. Some creatures of doubtful factuality are allowed into the 700 field when admitted into authorship, for example, spirits.[9]

Within the 6XX fields, a distinction is made between the 600 field ("Subject added entry—personal name") and the 650 field ("Subject added entry—topical heading"). The 600 field contains "Biblical characters except 'God,' the 'Devil,' angels and archangels, and gods such as 'Baal'; clans; families; and personal name subject headings with dates containing a phrase other than 'in fiction, drama, poetry, etc.'"[10]

The 650 field contains the Biblical characters excluded above, as well as many proper names, such as the Ellis Island example alluded to earlier in this article, fictional characters, and performing animals. One could argue that it has become a kind of "everything else" field.

Note that not only does the MARC format distinguish between the 7XX (descriptive cataloging) and 6XX (subject cataloging) fields, but it also distinguishes between X00 and X50 fields. X00 fields (600 and 700) are used only for real people (and their spirits). One question this research raises is whether or not it is important to exclude fictitious characters from X00 fields. If fictitious characters were placed in X00 fields, they could be indexed along with other personal names rather than with subjects. However, this solution would preclude double indexing as both name and subject and would fail to address the broader problem of the scattering of proper names between the two types of indexes.

The creation of a new tag in both the 6XX and the 7XX fields for proper names other than geographic, personal, or corporate would allow systems more freedom to index all proper names, including fictional characters. Those systems that call the indexes "name" and "subject" could put all proper names in the name index. Systems might also opt to provide both a name index and an author index: the first to accommodate users who are searching for names that are neither author nor subject, such as fictional characters, and the second to accommodate users who know that they are searching for an author in the narrow sense of the word. Another solution might be to double index some proper names in both the name/author and subject indexes. It would be important to create both a 6XX and a 7XX field to allow users to distinguish between works about the proper name and works associated with the proper name in some other way (e.g., works featuring the proper name as a character in fiction). It might be useful to have a different tag for performers than for authors in the narrow sense of writers of books. This would allow systems to put performers into broader "name" indexes rather than narrow "author" indexes.

Even if general indexes are widely adopted, the suggested change above to the MARC format would allow the design of system features to permit those users who know the MARC format to limit searches and specify more precisely the name and/or subject of interest.

REFERENCES AND NOTES

1. Joseph R. Matthews, *Public Access to Online Catalogs*, 2d ed. (New York: Neal-Schuman, 1985).
2. Library of Congress, Network Development and MARC Standards Office, *USMARC Format for Bibliographic Data, Including Guidelines for Content Designation* (Washington, D.C.: Library of Congress, 1988–).
3. James E. Ross, "Artists and Poets Online: Issues in Cataloging and Retrieval," *Cataloging & Classification Quarterly* 7, no.3:91–104 (Spring 1987); Thomas R. Sanders, "Uniform Title as Author: A Problem in Online Catalogs?" *Library Resources & Technical Services* 31, no.3:231–38 (July/Sept. 1987); Simone Klugman, "Failures in Subject Retrieval," *Cataloging & Classification Quarterly* 10, no.1/2:9–35 (1989); David M. Pilachowski and David Everett, "What's in a Name? Looking for People Online—Social Sciences," *Database* 8, no.3:47–65 (Aug. 1985).
4. *Guidelines on Subject Access to Individual Works of Fiction, Drama, Etc.* (Chicago: American Library Assn., 1990).
5. "Headings for Certain Entities," *Cataloging Service Bulletin* 43:2–11 (Winter 1989).
6. "Guidelines for Tagging Geographic Names," *Cataloging Service Bulletin* 47:60–61 (Winter 1990).
7. *USMARC Formats: Underlying Principles* (Chicago: Library and Information Technology Assn., 1989), rule 6.4.
8. Ben Tucker has pointed out to the authors that most performing animals' names are obscure and less well-known to the public than the names of the animal characters that they portray. For example, Lassie is the name of a fictitious character played by a number of different dogs with names of their own. These dogs are probably not widely known by their real names.
9. *Anglo-American Cataloguing Rules*, 2d ed., 1988 rev., (Ottawa: Canadian Library Assn.; Chicago: American Library Assn., 1988), 350, rule 21.26.
10. Library of Congress, Network Development and MARC Standards Office, *USMARC Format for Bibliographic Data*.

APPENDIX A. QUESTIONNAIRE

Does your library collect fiction, drama, films, or other materials with fictional characters?

_____ 1) Yes

_____ 2) No; if your answer is no, please do not continue with the questionnaire.

In answering the following questions, please make the following assumptions:

1. Access to individual works of fiction, drama, or other materials with fictional characters by means of fictional character names is widely available.
2. The library being used by the hypothetical user in the following questions does have the materials desired.

Please note: In the following questions, we have purposefully *not* defined terms such as 'name search' or 'author search.' We are interested in what users will *think* are retrievable in searches or indexes called by these names. In fact, definitions vary from system to system, and most users probably don't know our definitions.

Also please note: The pages of this questionnaire have questions on both sides!

1. A user would like to see what books featuring Sherlock Holmes as a principal character are available in the library.
   a) If an online public access catalog presented this user with the choice of a subject search vs. a name search, which do you think most users would choose *first* in order to search for Sherlock Holmes?
      _____1) Subject search
      _____2) Name search
   b) If an online public access catalog presented this user with the choice of subject search vs. an author search, which do you think most users would choose *first* in order to search for Sherlock Holmes?
      _____1) Subject search
      _____2) Author search

c) If a library patron looking for Sherlock Holmes stories were able to choose from the following categories, which do you think he or she would be most likely to choose? (Please choose just *one*.)

_____1) Author search
_____2) Name search
_____3) General search which does not require specifying the type of search at all
_____4) "Works about" search
_____5) Fictional character search
_____6) Subject search
_____7) Other (please write in): _____

2. A user would like to see what cartoons featuring Bugs Bunny as a principal character are available in the library.

a) If an online public access catalog presented this user with the choice of a name search vs. subject search, which do you think most users would choose *first* in order to search for Bugs Bunny?

_____1) Name search
_____2) Subject search

b) If an online public access catalog presented this user with the choice of an author search vs. a subject search, which do you think most users would choose *first* in order to search on Bugs Bunny?

_____1) Author search
_____2) Subject search

c) If a library patron looking for Bugs Bunny cartoons were able to choose from the following categories, which do you think he or she would be most likely to choose? (Please choose just *one*.)

_____1) Author search
_____2) Name search
_____3) General search which does not require specifying the type of search at all
_____4) "Works about" search
_____5) Fictional character search
_____6) Subject search
_____7) Other (please write in): _____

3. A user would like to see what books that discuss Bugs Bunny, give his history, etc., are available in the library.

a) If an online public access catalog presented this user with the choice of a subject search vs. name search, which do you think most users would choose *first* in order to search for Bugs Bunny?

_____1) Subject search
_____2) Name search

b) If an online public access catalog presented this user with the choice of a subject search vs. an author search, which do you think most users would choose *first* in order to search for Bugs Bunny?

_____1) Subject search
_____2) Author search

c) If a library patron looking for books about Bugs Bunny were able to choose from the following categories, which do you think he or she would be most likely to choose? (Please choose just *one*.)

_____1) Author search
_____2) Name search
_____3) General search which does not require specifying the type of search at all
_____4) "Works about" search
_____5) Fictional character search
_____6) Subject search
_____7) Other (please write in): _____

4. A user would like to see what works of fiction that feature Abraham Lincoln as a character are available in the library.

a) If an online public access catalog presented this user with the choice of a name search vs. subject search, which do you think most users would choose *first* in order to search for Abraham Lincoln?

_____1) Name search
_____2) Subject search

b) If an online public access catalog presented this user with the choice of an author search vs. a subject search, which do you think most users would choose *first* in order to search for Abraham Lincoln?
_____1) Author search
_____2) Subject search

c) If a library patron looking for books with Abraham Lincoln as a character were able to choose from the following categories, which do you think he or she would be most likely to choose? (Please choose just *one*.)
_____1) Author search
_____2) Name search
_____3) General search which does not require specifying the type of search at all
_____4) "Works about" search
_____5) Fictional character search
_____6) Subject search
_____7) Other (please write in): _____
_____

5. A user would like to see what biographies of Abraham Lincoln are available in the library.
a) If an online public access catalog presented this user with the choice of a subject search vs. a name search, which do you think most users would choose *first* in order to search for Abraham Lincoln?
_____1) Subject search
_____2) Name search

b) If an online public access catalog presented this user with the choice of an subject search vs. an author search, which do you think most users would choose *first* in order to search for Abraham Lincoln?
_____1) Subject search
_____2) Author search

c) If a library patron looking for a biography of Abraham Lincoln were able to choose from the following categories, which do you think he or she would be most likely to choose? (Please choose just *one*.)
_____1) Author search
_____2) Name search
_____3) General search which does not require specifying the type of search at all
_____4) "Works about" search
_____5) Fictional character search
_____6) Subject search
_____7) Other (please write in): _____
_____

6. In what type of library do you work?
_____1) public _____2) academic
_____3) other; please explain: _____
_____

7. How useful would it be for the patrons you serve to be able to search for individual works of fiction, drama, or other materials with fictional characters, by means of fictional character names?
_____1) Not at all useful _____2) Somewhat useful
_____1) Useful _____2) Very useful

8. What type(s) of catalog(s) does your library have? (Choose all that apply)
_____1) card _____2) microform _____3) online
_____4) other (please specify) _____

9. How frequently do you encounter users who seek individual works of fiction, drama, or other materials with fictional characters (whether they ask for them by fictional character name or not)?
_____1) 1 or more times a day
_____2) 1–4 times per week
_____3) less than once a week

10. How many years have you been employed as a reference librarian? _____
11. How many volumes does your library contain? _____
12. How many titles does your library contain? _____

If your library has an online catalog, please answer the following question. If not, you are finished!

Thank you very much for your help in this research.

13. Please answer the following two questions about your online catalog:

13a. Which of the following sets of searches most nearly approximates those available on your online catalog? (*Please check all that apply.* Also, we are not interested in the various title searches which may be available on your system.)

_____1) A search called author search, or a close equivalent, such as "A" or "AU."
_____2) A search called name search, or a close equivalent, such as "NA."
_____3) A search called subject search, or a close equivalent, such as "S" or "SU."
_____4) A general search that does require specifying the type of search at all (regardless of what it is called)
_____5) Other (please describe): _____

_____

13b. Which of the following searches would have to be used to find a name as a subject in your system, for example, a book about Abraham Lincoln? (Check as many as apply.)

_____1) Name search
_____2) Author search
_____3) Subject search
_____4) Not applicable (system does not require choice of a type of search).
_____5) Other (please explain): _____

_____

Thank you very much for your help in this research.

Performing Resource Usage Analysis for a NOTIS System

Mark Hinnebusch

Methods for determining the resources consumed by a NOTIS system are discussed. The metrics available for analysis are identified and monitoring record formats explained. Results of analysis of the NOTIS implementation at FCLA are presented.

The prevalent method of estimating resource needs for NOTIS systems appears to be anecdotal. This is not surprising given the effort required in the start-up phases of library automation projects and the relative youth of many NOTIS installations. The Florida Center for Library Automation (FCLA) has not had the luxury of avoiding the issue. Supporting nine state universities and two community colleges, FCLA topped 15 million transactions per month in October 1990, and it expects use to continue increasing. FCLA purchases computing time and services from the Northeast Regional Data Center (NERDC) of the State University System of Florida. As a result, budgets are driven by transaction cost and count estimates. This paper outlines the methods that the FCLA has developed to estimate both transaction costs and overall demand for NOTIS services. The results of these analyses are then presented.

METHODS OF TRANSACTION COST AND DEMAND ANALYSIS

NOTIS operates on IBM computers running either the DOS or MVS operating systems that use a transactional processing manager known as the Customer Information Control System (CICS). A NOTIS transaction is the work done by the computer between interactions with the terminal user. This is best shown by an illustration (see figure 1). Rarely is there one transaction per logical unit of work, be it an OPAC search, a circulation action, or a technical services activity. Nevertheless, the transaction, as the quantum of work in a CICS-based system, is the unit for which computing resource costs must be measured.

CICS manages system resources for NOTIS and it both monitors and reports use. The subsystem of CICS responsible for this is called the CICS Monitoring Facility (CMF). There are three types of monitoring data provided by CMF: accounting data, exception data, and performance data.

The accounting data are at the summary level for the system as a whole. They include the number of times any particular transaction code is invoked, and how many transactions any particular terminal or identifier invokes. Exception data are information identifying abnormal occurrences that may indicate potential problems. Performance data contain the resource usage metrics. These metrics are collectible on a transaction-by-transaction basis and for the system as a whole. Only the transactional level metrics are of interest, since global metrics can be derived during the analysis and broken out in greater detail than provided by global CMF reporting.

IBM provides the ability to collect monitoring information from their operating systems. This System Management Facility (SMF) provides the same types of information for the system as a whole as CMF does for CICS. CMF records are structured in such a way that they can be considered a special type of

Mark Hinnebusch is the Assistant Director for Technical Development at the Florida Center for Library Automation, Gainesville.

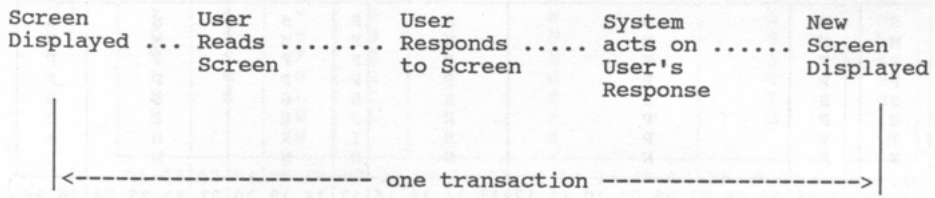


Figure 1. A CICS Transcation.

SMF record. A CMF record consists of an SMF header, an SMF product section, and a CICS section. In the event that CMF data are written to an SMF dataset, the SMF header and SMF product section make the CMF records look like other SMF records. Usually this is done for later analysis by the IBM-supplied GTFPARS program. A CMF record is defined as an SMF type 110 record. The CICS section is unique to CMF records. The CICS section always consists of a section header, a section descriptor, and section data. The first two bytes of the section header contain the length of the entire CICS section. Figures 2 through 4 show the major structures in a CMF record.

There are three types of CMF performance records: dictionary records, transaction performance records, and global performance records. All three record types share a common structure; however, the CICS section data areas have different contents for the different record types. In a dictionary record, the CICS section data area consists of multiple dictionary entries. Each dictionary entry defines a monitoring field that will appear in successive transaction and global performance records. Each dictionary entry contains a unique field number, the field name, the field format, the field length, the name of the generating CICS module, and a numeric field connector (see figure 5).

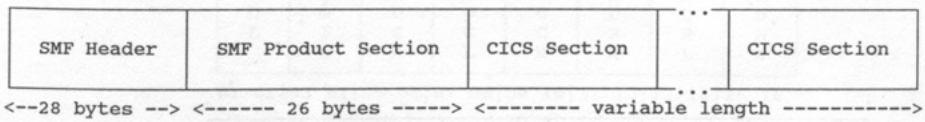


Figure 2. CMF Record Structure.

S	S	S	S	S	S	u	S	S	S
M	M	M	M	M	M	n	M	M	M
F	F	F	F	F	F	u	F	F	F
F	R	T	D	S	S	s	A	L	N
L	T	M	T	I	S	e	P	P	P
G	Y	E	E	D	S	d	S	S	S
01	02	03 04 05 06	07 08 09 10	11 12 13 14	15 16 17 18	19 20	21 22 23 24	25 26	27 28

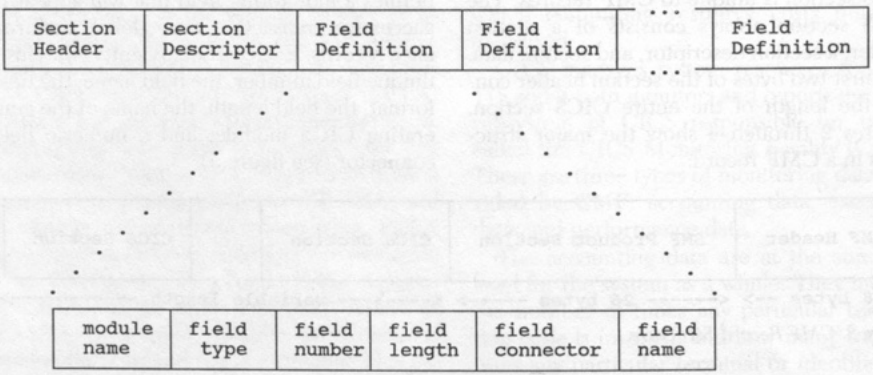
FIELDNAME	FORMAT	CONTENTS
SMFFLG	XL1	Operating System Indicator, always X'82'
SMFRTY	XL1	Record Type, always X'6E'
SMFTME	XL4	Time when the block was written in .01 seconds
SMFDTE	PL4	Date when the block was written as 00YYDDD+
SMFSID	CL4	System identifier
SMFSSS	CL4	Subsystem Identifier, indicates the CICS region
	XL2	Unused
SMFAPS	XL4	Offset of the product section from the block length prefix, always X'20'
SMFLPS	XL2	Length of product sections, always X'1A'
SMFNPS	XL2	Number of product sections, always 1

Figure 3. CMF SMF Header Structure.

SMFPSSTY	SMFPSRVR	SMFPSPRN	SMFPSRSN	SMFPSJID	SMFPSBKN	SMFPSLBW	SMFPSBAL
01 02	03 04	05 06 07 08	09 10 11 12	13 14 15 16	17	18 19 20	21 22 23 24	25 26

FIELDNAME	FORMAT	CONTENTS
SMFPSSTY	XL2	Product subtype, always X'0001' indicating monitoring
SMFPSRVR	CL2	Record version, always C'03'
SMFPSPRN	CL8	VTAM APPLID of CICS region
SMFPSRSN	PL4	Sequence number of block within CMF file
SMFPSJID	XL1	User journal ID for this CMF file
SMFPSBKN	XL3	Relative record number
SMFPSLBW	XL4	TTR of last record
SMFPSBAL	XL2	Number of bytes left in track

Figure 4. CMF SMF Product Section Structure.



Module name is 8 bytes, identifies the CICS module that produces the metric.
Field type is one byte.
Field number is a 3 character number.
Field length is a 2 byte hexidecimal value.
Field connector is a 2 byte hexidecimal value.
Field name is an 8 character string.

Figure 5. CMF Dictionary Record Section Structure.

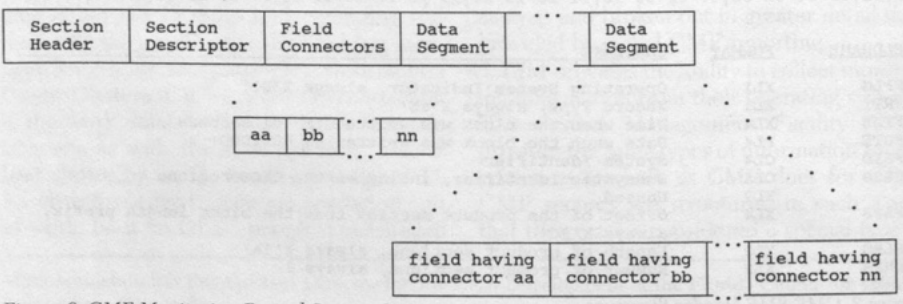
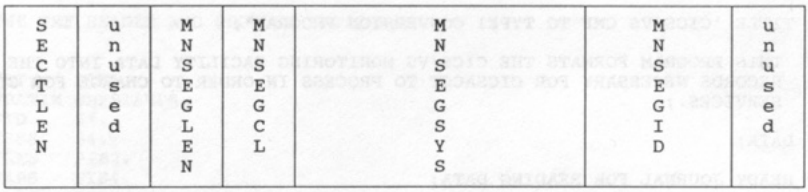
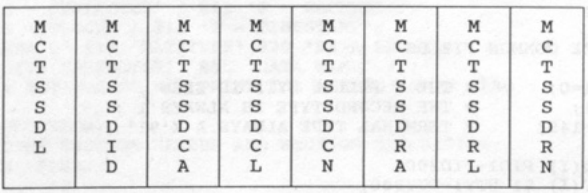


Figure 6. CMF Monitoring Record Section Structure.



FIELDNAME	FORMAT	CONTENTS
SECTLEN	XL2	Length of the CICS segment
	XL2	Unused, with SECTLEN makes up an RDW
MNSEGLEN	XL2	Length of the Section Header, excluding SECTLEN RDW, always X'12'
MNSEGCL	XL2	Class of data X'0002' - accounting X'0003' - performance X'0004' - exception
MNSEGSYS	CL8	VTAM APPLID of the CICS region
MNSEGID	CL4	Always 'CICS'
	XL2	Unused

Figure 7. CMF CICS Section Header Structure.



FIELDNAME	FORMAT	CONTENTS
MCTSSDL	XL2	Length of section descriptor and section data
MCTSSDID	XL2	Type of record X'0000' - dictionary X'0001' - transaction X'0002' - global performance
MCTSSDCA	XL2	Length of section descriptor, always X'0010'
MCTSSDCL	XL2	Length of each field connector, always X'0002'
MCTSSDCN	XL2	Number of connectors
MCTSSDRA	XL2	Offset of first data segment from start of section descriptor
MCTSSDRL	XL2	Length of each data segment
MCTSSDRN	XL2	Number of data segments

Figure 8. CMF CICS Section Descriptor Structure.

In transaction and performance records, the CICS section data area consists of multiple data segments (see figures 6 through 8.) Each data segment consists of multiple monitoring fields. A monitoring field in a transaction or global performance record is logically connected to its dictionary definition entry via the field connector. This connection is indirect in that the field connector for a particular field varies depending on the fields that the installation has chosen to monitor. The field number, on the other hand, is constant. Each transaction or performance record data section begins with a series of two-byte field connectors. There is one connector for each monitoring field in each of the succeeding

```
TITLE 'CICS/VS CMF TO TYPE1 CONVERSION PROGRAM';

*  THIS PROGRAM FORMATS THE CICS/VS MONITORING FACILITY DATA INTO THE TYPE 1
   RECORDS NECESSARY FOR CICSACCT TO PROCESS IN ORDER TO CHARGE FOR CICS/VS
   SERVICES.;

DATA;

* READY JOURNAL FOR READING DATA;

  INFILE JOURNAL LENGTH=BLOCKLEN COLUMN=COL STOPOVER END=ENDIN;
  FILE OUTPUTPR PRINT;

  READREC:
  INPUT SMFFLG PIB1. @;

*  CHECK TO SEE WHICH SYSTEM, IT MUST BE MVS;
   IF SMFFLG^=130 THEN DO;
      PUT // 120*'*';
      PUT 'RECORDS ARE NOT FROM AN MVS SYSTEM  (' SMFFLG= ')';
      PUT 'JOB ABORTED';
      PUT // 120*'*';
      ABORT 16;
      END;

*  CHECK TO SEE IF A RECORD IS JUST AN INITIALIZED JOURNAL RECORD.  ( RECORD
   LENGTH OF 42 BYTES) . IF IT IS, SKIP THE RECORD. ;

  IF BLOCKLEN<43 THEN DO;
       INPUT;
       GO TO READREC;
       END;

* INITIALIZE COMMON FIELDS . ;

   TRNRDWX1=0;       * THE 1 FILLER BYTE IS ZERO ;
   TRNRID=1;         * THE RECORD TYPE IS ALWAYS 1 ;
   TRNTRMT=145;      * TERMINAL TYPE ALWAYS A X'91' ;

  ARRAY PID(I) PID1-PID200;
  ARRAY PTY(I) $1 PTY1-PTY200;
  ARRAY PDL(I) PDL1-PDL200;
  ARRAY PHD(I) $40  PHD1-PHD200;

* THE RECORD DEFINITIONS THAT ARE USED ARE FROM THE FOLLOWING
  CICS/VS DSECTS:
          DFHJCSMF
          DFHCMPRC PREFIX
          DFHMCTSS
 ( THESE CAN BE OBTAINED FROM THE CICS/VS MACLIB ) ;
```

Figure 9. Listing of CMFPROG2.

data segments. Each data segment contains the same monitoring fields as every other data segment, and in all segments the order of the monitoring fields is the same and is congruent with the order of the field connectors. Each dictionary entry contains the field connector associated with the field that it is defining.

To find out how to process a data segment, one must inspect the field connectors at the head of the data section, use the field connector to look up the field definition in the dictionary, and use the contents of the dictionary entry to interpret the field. Since different monitoring fields are of different lengths, one must use the dictionary entries to control movement through the data segment (see figure 9 for an example of processing this data).

This convoluted process appears to have been designed so that different installations could choose to collect different subsets of monitoring data. Since each monitoring field is assigned a unique field identifier within its class (i.e., record type), the need for this level of redirection is not obvious. Nevertheless, IBM has chosen this methodology. Had they been aware of the MARC format and the NISO Z39.2 standard, they might have chosen to use it. The CMF use of the dictionary

```
*  READ THE SMF HEADER AND PRODUCT SECTION;

   INPUT
      SMFRTY    PIB1.
      SMFDATIM SMFSTAMP8.
      SMFID     $4.
      SMFSSS    $4.
      SMFRES    PIB2.
      SMFAPS    PIB4.
      SMFLPS    PIB2.
      SMFNPS    PIB2.
      SMFPSSTY PIB2.
      SMFPSRVR PIB2.
      SMFPSPRN $8.
      SMFPSRSN PIB4.
      SMFPSJID PIB1.
               +9    @;
   FORMAT SMFDATIM DATETIME. TRNTRMID $4. TRNOPN $4. TRNTRNID $4.;

   SMFMVS='MVS';

*  PUT THE SMF HEADER INFORMATION OUT FOR THE FIRST RECORD;
   IF SMFPRFLG^=1 THEN DO;
      PUT // 120*'*';
      PUT 'THE SYSTEM IS ' SMFMVS '. ' /
          'THE SMF RECORD TYPE IS ' SMFRTY  '. ' /
          'THIS RECORD WAS WRITTEN: ' SMFDATIM /
          'THE JOURNAL ID IS ' SMFPSJID '.' /
          'THE VTAM APPLICATION ID IS ' SMFPSPRN '.' ;
      PUT // 120*'*';
      PUT // 'DATATYPES:' / @12 'A - FULLWORD ' /
            @12 'C - CHARACTER' / @12 'P - PACKED' /
            @12 'S - CLOCK' / @12 'T - TIMESTAMP';
      PUT // @2 'DATAID' @10 'DATATYPE' @20 'DATA LENGTH'
            @35 'DATA CONNECTOR'  @55 'DATA NAME' //;
      SMFPRFLG=1; * SET THE DO NOT PRINT THE RECORD AGAIN FLAG ;
      END;
   DO WHILE (COL<=BLOCKLEN);
      *  READ THE CICS SECTION HEADER AND SECTION DESCRIPTOR;
      INPUT SECTLEN  PIB2.
                     +2
            MNSEGLEN PIB2.
            MNSEGCL  PIB2.
            MNSEGSYS $8.
            MNSEGID  $4.
                     +2
            MCTSSDL  PIB2.
            MCTSSDID PIB2.
            MCTSSDCA PIB2.
            MCTSSDCL PIB2.
            MCTSSDCN PIB2.
            MCTSSDRA PIB2.
            MCTSSDRL PIB2.
            MCTSSDRN PIB2. @;
```

Figure 9. Listing of CMFPROG2 (cont.).

record offers the possibility of dynamic dictionaries. The dictionary record would be analogous to encoding a tag table in a MARC record and transmitting it along with the records that it defines.

CMF transaction level records may contain any of the metrics listed in appendix A. Which metrics are actually collected is an installation option. At NERDC, the complex CMF records are reduced and reformatted nightly to fixed length, fixed-position records containing the metrics used by the charging algorithm. For historical reasons, the reduced records are called "Type 1" records. The format of these records is shown in table 1. The program (CMFPROG2) used to convert from CMF format to the "Type 1" flat record is listed in figure 9.

IBM offers the CICSPARS program to produce batch reports based on CMF data. It also operates interactively, allowing online display of CMF data at the summary level. Because of an interest in associating detailed performance data with NOTIS functions, the

```
*  CHECK FOR A DICTIONARY RECORD.  IF THIS IS NOT
   A DICTIONARY RECORD THEN IT IS A MONITORING DATA RECORD.  ;

   IF  MCTSSDID=0 THEN DO;

    * THIS IS A DICTIONARY RECORD.  CREATE THE VARIABLES FROM THE
      DICTIONARY SO THE DATA RECORDS CAN BE READ AND FORMATTED
      EITHER FOR OUTPUT LISTINGS OR TO CREATE THE TYPE1 RECORD. ;

    * THIS SHOULD BE A PERFORMANCE CLASS RECORD, IF NOT, THEN SKIP ;

      LENGTH CMFCLASS $ 20;
      IF MNSEGCL^=3 THEN DO;
                    INPUT;
                    RETAIN;
                    RETURN;
                    END;

      CMFCLASS='PERFORMANCE';

    * LOOP THROUGH THE DICTIONARY RECORD, READING THE FIELDS WHICH
      DEFINE THE DICTIONARY AND BUILDING ARRAYS. ;

      DO WHILE(MCTSSDRN^=0)    /* NUMBER OF DATA SEGMENTS>0 */;
          INPUT  DATAMODL $8.
                 DATATYPE $1.
                 DATAID   $3.
                 DATAL    PIB2.
                 DATACN   PIB2.
                 DATANM   $8. @;
          I=DATACN;
          IF I<1 OR I>200 THEN DO;
            PUT  'ERROR: DICTIONARY ENTRY OUT OR RANGE' / MCTSSDRN= /;
            PUT  DATAMODL $8.
                 DATATYPE $1.
                 DATAID   $3.
                 DATAL    PIB2.
                 DATACN   PIB2.
                 DATANM   $8. @;
            END /* BAD DICTIONARY ENTRY */;

        * ADD TO THE PERFORMANCE CLASS DICTIONARY ARRAY. ;

          ELSE DO;
            PID=DATAID;
            PTY=DATATYPE;
            PDL=DATAL;
            PHD=DATANM;
```

Figure 9. Listing of CMFPROG2 (cont.).

FCLA developed its own suite of customized reporting programs. The SAS program that reports resource usage, LIX15, is reprinted in figure 10. Output from the program for November 1988 is shown in figures 11A to 11J.

CICS operator identifiers (OPERIDs) are defined in our implementation such that the first character identifies the user's institution, the second character defines the processing unit, and the third character defines the functional area. This breakdown is shown in appendix B. Public access terminals within the libraries are locked into LUIS (the public online catalog) using a unique CICS OPERID, #$#. Using the CICS transaction codes, which identify the database being used, and the OPERIDs, it is possible to associate most of the resource records with a particular institution and functional area.

RESULTS OF ANALYSIS OF THE FCLA NOTIS SYSTEM

One of the initial problems to be resolved in performing usage analysis on a heavily used system is to define an appropriate sample of the data. In fiscal year 1987–88, from July 1, 1987, through June 30, 1988, FCLA had more than 77 million transactions, of which roughly

```
      * OUTPUT EACH DICTIONARY IDENTIFIER SO THAT
        WE HAVE A HARDCOPY LISTING WHICH DESCRIBES THE
        FIELDS OF THE RECORDS WE ARE WORKING WITH.
        ;

        PUT DATAID @15 DATATYPE @22 DATAL @40 DATACN
            @55 DATANM $8. ;
        MCTSSDRN=MCTSSDRN-1;
        END /* GOOD DICTIONARY ENTRY */;

      END  /*  DO WHILE(MCTSSDRN^=0) (DATA SEGMENT)  */;

    END /* IF MCTSSDID=0 (DICTIONARY PROCESSING) */;


*   THIS BEGINS THE CODE IF THE RECORD IS NOT A DICTIONARY
    RECORD.  INPUT THE INDEX AT THE BEGINNING OF EACH
    RECORD.  THIS TELLS WHAT IDENTIFIERS IN THE DICTIONARY
    ARE CONTAINED IN THIS RECORD.  USE EACH BYTE OF THIS
    INDEX TO INDEX INTO THE APPROPRIATE ARRAY (I.E., ACCOUNT,
    PERFORMANCE, OR EXCEPTION).
    ;


  ELSE DO;

    CON_LEN=MCTSSDCN*2;  /* 2 BYTES PER CONNECTOR */;
    INPUT INDEX $VARYING100. CON_LEN @;
    OFFSET=1;
    INCNT=INCNT+1;

  * IT IS POSSIBLE FOR A PHYSICAL RECORD TO CONTAIN MORE THAN
    ONE DATA RECORD. HERE ARE TWO LOOPS, THE FIRST LOOP
    KEEPS COUNT OF THE NUMBER OF DATA RECORDS IN A GIVEN
    PHYSICAL RECORD.  THE SECOND LOOP READS EACH FIELD OF THE
    DATA RECORD AND OUTPUTS EITHER  A LISTING OR CREATES THE
    APPROPRIATE TYPE 1 RECORD FROM THE DATA.
    ;

    COUNTER1=MCTSSDRN;
    COUNTER2=MCTSSDCN;
    DO WHILE(COUNTER1^=0);
       DO WHILE(COUNTER2^=0);

         INDEXC=SUBSTR(INDEX,OFFSET,2);
         I=INPUT(INDEXC,PIB2.);

        * CHECK FOR PERFORMANCE DATA RECORD, IF NOT ABORT. ;
         IF MNSEGCL^=3 THEN DO;
            PUT // 120*'*' /  'THERE ARE NO PERFORMANCE RECORDS';
            PUT 'THIS JOB ABORTED' /// 120*'*';
            ABORT 16;
            END;
```

Figure 9. Listing of CMFPROG2 (cont.).

50 million were accessing LUIS, the online public catalog. The cost and logistics of using the full set of resource records make an exhaustive analysis impossible. Use of the system throughout the academic year is cyclic, as seen in figure 12. While it seemed improbable that there would be month-to-month variations in per-transaction resource use or in the mix of searches and resource use per search, an attempt was made to choose an "average" month. November 1988 was chosen, since it was neither a peak nor a valley month and was the most recent data available at the time of the analysis. FCLA is not currently running the keyword and Boolean search feature of NOTIS, so the analysis does not include this search type.

It must be noted that the results of the

```
            * WE USE THE PERFORMANCE DATA RECORD FOR COLLECTING THE
              DATA NECESSARY TO CREATE THE TYPE1 RECORD FOR THE
              CICSACCT PROGRAM TO READ IN ORDER TO DO THE CHARGING.;

              FIELDL=PDL;
              INPUT FIELD $VARYING8. FIELDL @;
              IF (PTY='A' AND PDL=1) THEN
                          PDATA=INPUT(FIELD,PIB1.);
                 ELSE IF (PTY='A' AND PDL=2) THEN
                          PDATA=INPUT(FIELD,PIB2.);
                 ELSE IF (PTY='A' AND PDL=4) THEN
                                    PDATA=INPUT(FIELD,PIB4.);
                 ELSE IF (PTY='A' AND PDL=6) THEN
                          PDATA=INPUT(FIELD,PIB6.);
                 ELSE IF (PTY='A' AND PDL=8) THEN
                          PDATA=INPUT(FIELD,PIB8.);
                 ELSE IF (PTY='P' AND PDL=2) THEN
                          PDATA=INPUT(FIELD,PD2.);
                 ELSE IF (PTY='P' AND PDL=3) THEN
                          PDATA=INPUT(FIELD,PD3.);
                 ELSE IF (PTY='P' AND PDL=4) THEN
                          PDATA=INPUT(FIELD,PD4.);
                 ELSE IF (PTY='T' AND PDL=8) THEN
                          PDATA=INPUT(FIELD,TODSTAMP.);
                 ELSE IF (PTY='S' AND(PDL=4 OR PDL=8)) THEN DO;
                          PDATA=INPUT(FIELD,PIB4.);
                          PDATA=PDATA*.000016;
                          END;

IF _ERROR_=1 THEN DO;
IF ERRCNT<50 THEN PUT 'ERROR ENCOUNTERED:' I= PTY= PDL= PDATA= FIELD=;
  ERRCNT=ERRCNT+1;
  _ERROR_=0;
  END;

                 LINK CICSACCT;
                 OFFSET=OFFSET+2;
                 COUNTER2=COUNTER2-1;
                 END  /* DO WHILE(COUNTER2^=0) */;

           LINK WRITEREC;
           OFFSET=1;
           COUNTER2=MCTSSDCN;
           COUNTER1=COUNTER1-1;
           END   /*  DO WHILE(COUNTER1^=0) */;

      END   /* ELSE (NOT DICTIONARY PROCESSING) */;

   END   /* BLOCK */;
   INPUT;
   RETAIN;
   RETURN;
```

Figure 9. Listing of CMFPROG2 (cont.).

FCLA analysis methods can be generalized only by recognizing that differences in hardware and software configurations as well as installation parameter choices may have an impact on transaction resource mix. Variations in LUIS training styles at different institutions may affect LUIS searching. Other NOTIS sites are encouraged to use the programs printed herein and to report their results.

TRANSACTION ANALYSIS RESULTS

Summary transaction resource usage analysis was performed on all NOTIS transactions as a group. Then, the statistics were calculated for each institution group, an institution group being either an actual university or a special function. This was done to ascertain the effects on transaction resource demands of database size and demand. Transactions were also grouped as LUIS, circulation, or technical services to ascertain the basic resource costs for these services. The average resource costs per transaction were plotted as a function of hour of day to ascertain system load effects

```
*****************************************************************
*           SUBROUTINES                                         *
*****************************************************************;

CICSACCT:
   *  THIS IS THE ROUTINE LINKED TO BY THE PERFORMANCE RECORD SECTION OF THE
   MAIN PROGRAM.  HERE THE ACCOUNTING DATA IS PULLED FROM THE PERFORMANCE
   CLASS RECORD OF CMF AND, IF NEEDED, CONVERTED TO THE APPROPRIATE TYPES
   AND/OR UNITS THAT MAKE UP THE TYPE 1 RECORD FOR CICSACCT PROGRAM TO USE FOR
   CHARGING.

   CHECK THE DICTIONARY IDENTIFIER FOR THE DATA FIELD.  ASSIGN THAT FIELD TO
   THE COMPARABLE FIELD IN THE TYPE 1 RECORD.
  ;

    IF PID=1 THEN TRNTRNID=FIELD;     * TRANSACTION ID;
    ELSE IF PID=2 THEN TRNTRMID=FIELD; * TERMINAL ID;
    ELSE IF PID=3 THEN TRNOPN=FIELD; * OPERATOR ID;
    ELSE IF PID=5 THEN DO;
         TRNITIME=TIMEPART(PDATA)*100;       * INPUT TIME ;
         TRNDATE=JULDATE(DATEPART(PDATA));  * GET DATE;
         END;
    ELSE IF PID=6 THEN TRNTIME=TIMEPART(PDATA)*100;
         * TRNTIME IS DETACH TIME (STOP TIME) ;
         * TRNTIME IS THE TIME THE RECORD WAS WRITTEN ;
    ELSE IF PID=8 THEN DO;
         PDATA=PDATA/.000026045; * CONVERT TO TIMER UNITS ;
         TRNRTIME=PDATA;
         END;

*  THE NEXT FIVE CALCULATE THE RESPONSE TIME AS THE SUM OF CPU TIME,
   TASK DISPATCH TIME, WAIT FOR TC I/O, WAIT FOR FC I/O,
   WAIT FOR JC I/O, WAIT FOR TS I/O AND TIME ON SUSPEND CHAIN.
  ;
    ELSE IF PID=9 THEN TRNRTIME=TRNRTIME+PDATA; * WAIT FOR TC I/O;
    ELSE IF PID=10 THEN TRNRTIME=TRNRTIME+PDATA; * WAIT FOR JC I/O;
    ELSE IF PID=11 THEN TRNRTIME=TRNRTIME+PDATA; * WAIT FOR TS I/O;
    ELSE IF PID=14 THEN TRNRTIME=TRNRTIME+PDATA; * TIME ON SUSPEND;
    ELSE IF PID=63 THEN TRNRTIME=TRNRTIME+PDATA; * WAIT FOR FC I/O;

    ELSE IF PID=33 THEN TRNMAXC=PDATA; * CHECK FOR MAXIMUM CORE ;
    ELSE IF PID=70 THEN TRNDIO=PDATA;

*   CONVERT THE CORE RESIDENCE WHICH IS IN  UNITS OF
    BYTE/16 MICROSECONDS TO 64K BYTE MILLISECONDS;

    ELSE IF PID=73 THEN TRNRES=(PDATA/65536)*500;
    ELSE IF PID=83 THEN TRNTIO=TRNTIO+PDATA; * TERMINAL I/O COUNTER ;
    ELSE IF PID=84 THEN TRNTIO=TRNTIO+PDATA; * TERMINAL I/O COUNTER ;
    ELSE IF PID=85 THEN TRNTIO=TRNTIO+PDATA; * TERMINAL I/O COUNTER ;
    ELSE IF PID=86 THEN TRNTIO=TRNTIO+PDATA; * TERMINAL I/O COUNTER ;
    ELSE IF PID=28 THEN TRNCPU=PDATA; * CPU SECONDS ;
    RETURN;
```

Figure 9. Listing of CMFPROG2 (cont.).

on per transaction resource consumption.

Inspection of the resource usage statistics for the system as a whole provides the following information (see figure 11A).

Disk I/O (DISKIO)

There is a huge variation between the mean (16) and the maximum (10,455) number of disk I/O operations per transaction. This is probably due to a small number of titles with extremely large numbers of associated item records. Whether or not this situation drives excessive costs depends on the number of transactions with large disk I/O demands (i.e., the dispersion of the disk I/O per transaction metric).

CPU Usage (CPUSEC)

The disparity between the maximum and mean value merits attention. This variation may be due to the same situation as disk I/O variation.

```
WRITEREC:

*  THIS ROUTINE CHANGES THE FILE OUTPUT TO THE DATASET THAT WILL
   CONTAIN THE TYPE 1 RECORDS.
   IT CHANGES THE FILE TO THE DD CARD NAMED OUTPUT, WRITES THE
   RECORD AND THEN CHANGES THE FILE BACK TO THE PRINT DATASET.

   CHECK TO SEE IF THE RECORD IS A SYSTEM TASK RECORD.
   IF IT IS, DO NOT OUTPUT TO THE TYPE1 DATASET.
   ;

   IF TRNTRNID=:'KC.' THEN RETURN;
   ELSE IF TRNTRNID=:'CCMF' THEN RETURN;
   FILE OUTPUT;
   TRNRTIME=TRNRTIME*100;
   TRNSMFID=SMFID;
   TRNSSSID=SMFSSS;
   TRNAPPLI=SMFPSPRN;
   PUT
        TRNRDWX1 PIB1.
        TRNRID   PIB1.
        TRNTIME  IB4.
        TRNDATE  PD4.
        TRNTRMID $4.
        TRNOPN   $4.
        TRNTRNID $4.
        TRNITIME PIB4.
        TRNRTIME PIB4.
        TRNCPU   PIB4.
        TRNRES   PIB4.
        TRNTIO   PIB4.
        TRNMAXC  PIB4.
        TRNDIO   PIB2.
        TRNSMFID $4.
        TRNSSSID $4.
        TRNAPPLI $8.
          ;
   OUTCNT+1;    RETAIN OUTCNT;
   FILE OUTPUTPR PRINT;
   RETURN;
```

Figure 9. Listing of CMFPROG2 (cont.).

Table 1. NERDC "Type 1" Record Format.

Position	Contents
01-01	Reserved
02-02	Record type indicator, always X'01'
03-06	Time that CMF record was written
07-10	Date of transaction, as Packed Decimal Julian
11-14	Terminal ID
15-18	Operator ID
19-22	Transaction ID
23-26	Time transaction started, in hundredths of a second since midnight
27-30	Response time, in hundredths of a second
31-34	CPU use, in clock ticks; 38,400 per second.
35-38	Core residence, in byte-seconds.
39-42	Terminal I/O; sum of bytes input and output.
43-46	Maximum storage, i.e. high-water mark, in bytes.
47-48	Disk I/O; Number of FCT calls times 3.
49-52	System Identifier
53-56	Subsystem Identifier, i.e. CICS region ID
57-64	VTAM APPLID

```
OPTIONS PAGESIZE=55;
TITLE 'FLORIDA CENTER FOR LIBRARY AUTOMATION';
TITLE4 'RESOURCE USAGE ANALYSIS';

*       INPUT RESOURCE USAGE RECORDS (SMF TYPE 1), WHERE THERE
*       IS ONE RECORD PER TRANSACTION, WHERE A TRANSACTION IS
*       DELIMITED BY USE OF THE ENTER KEY, A FUNCTION KEY, A PA
*       KEY, OR THE ATTN KEY.  RECORDS ARE WRITTEN BY CMFPROG2;

DATA USAGE.USAGE; INFILE SMF1;
     INPUT  +1
            TYPE IB1.
            +4
            DATETRM PD4.
            TERMINAL $4.
            +3
            +1
            TRANID $4.
            TMETRAN IB4.
            RESPON IB4.
            CPU IB4.
            CORE IB4.
            TERMIO IB4.
            MAXSTOR IB4.
            DISKIO IB2. ;

*            KEEP ONLY LIBRARY TRANSACTION TYPE 1 RECORDS;

     IF TYPE ^= 1 THEN DELETE;
     IF SUBSTR(TRANID,1,2) ^= 'LU'
     AND SUBSTR(TRANID,1,2) ^= 'L#'
     AND SUBSTR(TRANID,1,2) ^= 'L$'
     AND SUBSTR(TRANID,1,2) ^= 'LT'
     AND SUBSTR(TRANID,1,2) ^= 'LV'
     AND SUBSTR(TRANID,1,2) ^= 'LX'
     AND SUBSTR(TRANID,1,2) ^= 'LC'
     AND SUBSTR(TRANID,1,2) ^= 'LP' THEN DELETE;

*           SET FUNCTION AND INSTITUTION BASED ON TRANID;

FUNCTION=REPEAT(' ',39);
INST=REPEAT(' ',39);

     IF SUBSTR(TRANID,1,2)='LU' OR SUBSTR(TRANID,1,2)='L#' THEN DO;
         FUNCTION='LUIS          ';
         INST=SUBSTR(TRANID,3,2);
         END;
```

Figure 10. Listing of LIX15.

Terminal I/O (TERMIO)

The terminal I/O statistics present no surprises as they are within reasonable and expected limits.

Maximum Storage (MAXSTOR)

The highwater mark for storage ranges between 6,960 bytes and 50,656 bytes—an acceptable range.

Core Residency (TWOKMINS)

Core residence, measured as minutes of 2 kilobytes of memory use, reflects both transaction duration and memory use. Given the CPU usage variation, it seems reasonable to ascribe variation in core residency to the same cause as variation in CPU usage.

Response Time (RESPONSE)

The measured response time is at the CPU, not at the terminal. Telecommunications overhead is not reflected. There is a great deal of variation in response time, yet response times are so short that they are negligible.

Figure 11B shows the same analysis performed for each institution supported by FCLA as well as some specialized services. In the following discussion, many items are excluded. This includes the database selection menu, because it is a completely separate, non-NOTIS software system that exists only at FCLA. The FCLA database, consisting pri-

```
    IF SUBSTR(TRANID,1,2) = 'LT' THEN DO;
        FUNCTION='TECH SERVICES';
        INST=SUBSTR(TRANID,3,2);
        IF INST='HC' THEN INST='UF';
        IF INST='LL' THEN INST='UF';
        IF INST='TL' THEN INST='UF';
        IF INST='ML' THEN INST='UF';
        IF INST='JH' THEN INST='UF';
        IF INST='FL' THEN INST='FS';
        END;

    IF SUBSTR(TRANID,1,2) = 'LV' THEN DO;
        FUNCTION='MANAGEMENT    ';
        INST='FCLA OVERHEAD';
        END;

    IF SUBSTR(TRANID,1,2) = 'LX' THEN DO;
        FUNCTION='MANAGEMENT    ';
        INST='FCLA OVERHEAD';
        END;

    IF SUBSTR(TRANID,1,2) = 'LC' OR
       SUBSTR(TRANID,1,2) = 'L$' OR
       SUBSTR(TRANID,1,2) = 'LP' THEN DO;
        FUNCTION='CIRCULATION   ';
        IF SUBSTR(TRANID,3,1)='A' THEN INST='UF';
        IF SUBSTR(TRANID,3,1)='B' THEN INST='CF';
        IF SUBSTR(TRANID,3,1)='C' THEN INST='NF';
        IF SUBSTR(TRANID,3,1)='D' THEN INST='WF';
        IF SUBSTR(TRANID,3,1)='E' THEN INST='SF';
        IF SUBSTR(TRANID,3,1)='F' THEN INST='FS';
        IF SUBSTR(TRANID,3,1)='G' THEN INST='FA';
        IF SUBSTR(TRANID,3,1)='H' THEN INST='FI';
        IF SUBSTR(TRANID,3,1)='I' THEN INST='AM';
        IF SUBSTR(TRANID,3,1)='J' THEN INST='C1';
        IF SUBSTR(TRANID,3,1)='X' THEN INST='SP';
        IF SUBSTR(TRANID,3,1)='Y' THEN INST='TF';
        IF SUBSTR(TRANID,3,1)='Z' THEN INST='FC';
        IF SUBSTR(TRANID,3,1)='0' THEN INST='TE';
        END;

*           BUILD DATE STRING;

    DATE=DATEJUL(DATETRM);
    FORMAT DATE WEEKDATE. ;

*           ADJUST DATE TO TV GUIDE TIME;

  TIME=TMETRAN;
  HOUR=HOUR(TIME);
  IF HOUR<6 THEN DATE=DATE-1;
  MONTH=MONTH(DATE);
```

Figure 10. Listing of LIX15 (cont.).

marily of records defining the IBM manual collection in the FCLA office, does not represent the average library collection and is therefore excluded. The newspaper index has no holdings attached, so it is excluded. Because the SUS extension library has essentially no LUIS use—it serves the other university libraries directly and has no patrons per se—it is also excluded. The union author/title index is an entry point only, and the resources consumed after the initial index selection would be affected by the particular records selected, so the union index is also excluded from the analysis. These exclusions reduce the count of transactions in the analysis from 9,864,412 to 9,577,563, a reduction of 2.91 percent.

Calculations were performed to ascertain the effect of institution size on per-transaction resource demands. The mean is the appropriate statistic to use in studying this question. A quick way of identifying potentially significant variation is to look at the range for each metric across the various institutions and ascertain what percentage of the minimum

```
*           CALCULATE CPU CHARGE;

     CPUSEC=(CPU/38400);
     CPUCHG=CPUSEC*0.44;
     CPUMIN=CPUSEC/60;

*           CALCULATE CORE RESIDENCE CHARGE;

     TWOKMINS=CORE/(32*500*120);
     CORECHG=TWOKMINS*0.089;

*           CALCULATE TERMINAL I/O CHARGES;

     TERMIO=(TERMIO/2048);
     TRMIOCHG=TERMIO*0.01;

*           CALCULATE DISK I/O CHARGES;

     EXCPCHG=(DISKIO/1000)*0.20;

*           CALCULATE TOTAL CHARGE, INIT TIME, AND RESPONSE TIME;

     TOTCOST=CPUCHG+CORECHG+EXCPCHG+TRMIOCHG;

*           TMETRAN IS IN .01 SECONDS SINCE MIDNIGHT;
*           INITTIME IS IN SECONDS SINCE MIDNIGHT;

     INITTIME=FLOOR(TMETRAN/100);
     RESPONSE=RESPON/100;
     HOUR=FLOOR(INITTIME/3600);

PROC SORT DATA=USAGE.USAGE OUT=SRTUSAGE.USAGE;
     BY INST;

PROC MEANS;
     VAR TERMIO MAXSTOR DISKIO CPUSEC TWOKMINS RESPONSE;

PROC MEANS;
     BY INST;
     VAR TERMIO MAXSTOR DISKIO CPUSEC TWOKMINS RESPONSE;


PROC SORT DATA=USAGE.USAGE OUT=SRTUSAGE.USAGE;
     BY FUNCTION;

PROC MEANS;
     BY FUNCTION;
     VAR TERMIO MAXSTOR DISKIO CPUSEC TWOKMINS RESPONSE;


PROC SORT DATA=USAGE.USAGE OUT=SRTUSAGE.USAGE;
     BY HOUR;
```

Figure 10. Listing of LIX15 (cont.).

the range is for any institution. The results of these calculations are shown in table 2. The variable identifying the resource in the computer runs is found below the resource. The institutions responsible for each mean value are shown in parentheses under each value. The abbreviations for the institutions are found in appendix B.

From table 2, it appears that institution size may be a factor, since SFCC is the smallest institution served by FCLA and UF is the largest. To test this, collection size was used as a measure of institution size. The mean disk I/O and mean core residence per 1,000 titles were calculated. Table 3 shows the distribution of mean disk I/Os for each of the institutions. Table 4 shows the distribution of mean core residence across institutions. From this test, there appears to be no correlation between the size of the institution and the average number of disk I/Os and core residency. In fact, these metrics appear to be fairly independent of collection size.

Transactions were also broken down by func-

```
PROC MEANS NOPRINT;
     BY HOUR;
     OUTPUT OUT=B
     SUM= TOTCPU TOTDISK TOT2KMIN TOTRESP
     MEAN= AVGCPU AVGDISK AVG2KMIN AVGRESP;
     VAR CPUSEC DISKIO TWOKMINS RESPONSE;


PROC PLOT;
     PLOT TOTCPU*HOUR;
     PLOT TOTDISK*HOUR;
     PLOT TOT2KMIN*HOUR;
     PLOT AVGCPU*DATE;
     PLOT AVGDISK*DATE;
     PLOT AVG2KMIN*DATE;
     PLOT AVGRESP*DATE;
```

Figure 10. Listing of LIX15 (cont.).

Table 2. Effect of Institution Size on Resource Demands (Preliminary Test).

Resource	Low Mean Usage	High Mean Usage	Mean Range	% Mean Is of Lowest Mean
Terminal I/O	.32983	.41844	.08861	26.68
TERMIO	(SFCC)	(USF)		
Max Storage	11065	12000	935	8.45
MAXSTOR	(FAU)	(UF)		
Disk I/O	10.98780	21.24255	10.25475	93.33
DISKIO	(SFCC)	(UF)		
CPU Use	.01543	.01753	.0021	13.61
CPUSEC	(SFCC)	(UF)		
Core Residence	.02229	.04996	.02767	124.14
TWOKMINS	(SFCC)	(UF)		
Response Time	.10201	.12270	.02069	20.28
RESPONSE	(FSU)	(USF)		

Table 3. Mean Disk I/O as a Function of Institution Collection Size.

Institution Titles	No. of Titles	Mean Disk I/O	Mean Disk I/O per 1,000
Santa Fe Community College	52,662	11.0	0.2088
Florida A&M University	263,699	12.6	0.0478
University of North Florida	310,112	13.6	0.0438
University of West Florida	300,772	14.3	0.0475
University of Central Florida	328,249	13.0	0.0396
University of South Florida	945,303	15.1	0.0159
University of Florida	1,122,612	21.2	0.0188
Florida Atlantic University	357,692	13.2	0.0369
Florida International University	388,483	13.6	0.0350
Florida State University	1,125,533	14.7	0.0130
Total system	5,195,117	16.4	0.0031

Table 4. Mean Core Residence as a Function of Institution Collection Size.

Institution Titles	No. of Titles	Mean Core Residence	Mean Core Residence per 1,000
Santa Fe Community College	52,662	.0223	4.234×10^{-4}
Florida A&M University	263,699	.0251	0.951×10^{-4}
University of North Florida	310,112	.0273	0.880×10^{-4}
University of West Florida	300,772	.0293	0.974×10^{-4}
University of Central Florida	328,249	.0255	0.776×10^{-4}
University of South Florida	945,303	.0312	0.330×10^{-4}
University of Florida	1,122,612	.0500	0.445×10^{-4}
Florida Atlantic University	357,692	.0258	0.721×10^{-4}
Florida International University	388,483	.0274	0.705×10^{-4}
Florida State South Florida	1,125,533	.0298	0.264×10^{-4}
Total System	5,195,117	.0354	0.068×10^{-4}

tion. The results are shown in figure 11C. These results show that circulation consumes more resources than LUIS or technical services and has the longest response time. However, circulation makes up only 13 percent of the total system use. Technical services, which makes up 20 percent of system use, consumes the least amount of system resources.

From the "Maximum Value" column in figure 11C, note that the transactions consuming extravagant amounts of resources are LUIS transactions. Isolating these transactions, FCLA found that they are extremely rare. This is due to titles with large numbers of associated items such as the New York Times, of which the University of Florida has thousands of item records. Such evidence argues for a close look at how LUIS processes these types of titles.

Figures 11D through 11F show total consumption of three resources as a function of time of day. CPU seconds, disk I/Os, and core residence all share the same shape and are all driven by actual transaction counts per hour. System use peaks between 10 and 11 a.m. and then again between 3 and 4 p.m. A smaller peak occurs at 8 to 9 in the evening.

Figures 11G through 11J show average per-transaction resource usage as a function of hour of day. Average CPU usage spikes in the early morning. This may be due to transactions that automatically lock terminals into LUIS. After this peak, CPU usage stabilizes for the rest of the day. Average disk I/O appears to climb throughout the day, which seems remarkable at first glance. However, the shape of the graph seems to reflect both the mix of technical services, circulation, and LUIS transactions throughout the day and the differences in mean disk I/O activity for the different transaction types. Notice that the first peak is between noon and 1 p.m., when most technical services personnel are at lunch, and LUIS usage as a percent of overall system use increases sharply. The next peak is between 5 and 6 p.m. when staff use drops off significantly. Another peak is between 10 and 11 p.m. when circulation as a percentage of system use increases dramatically. The last circulation peak is after midnight. All of our libraries close at 1 a.m. Eastern Time, so this peak is expected. A further check would be to plot resource usage by time of day for each type of transaction.

Core residence follows the same general shape as disk I/O but the variation is less extreme. This is expected, since the difference between LUIS mean core residence and technical services mean core residence is not as extreme as is the difference in disk I/O.

VARIABLE	N	MEAN	STANDARD DEVIATION	MINIMUM VALUE	MAXIMUM VALUE	STD ERROR OF MEAN	SUM
TERMIO	9864412	0.40865	0.18641	0.00000	1.37939	0.00009272	4031091
MAXSTOR	9864412	11644.36840	3051.96158	6960.00000	50656.00000	1.51799288	114861216617
DISKIO	9864412	16.38941	80.44109	0.00000	10455.00000	0.04001000	161671894
CPUSEC	9864412	0.01651	0.01218	0.01000	2.33945	0.00000606	162861
TWOKMINS	9864412	0.03539	0.23457	0.00066	52.31238	0.00011667	349101
RESPONSE	9864412	0.11390	0.11138	0.02000	13.49000	0.00005540	1123556

Figure 11A: LIX15 Output for November 1988: All Transactions.

VARIABLE	N	MEAN	STANDARD DEVIATION	MINIMUM VALUE	MAXIMUM VALUE	STD ERROR OF MEAN	SUM
---------- INST=FLORIDA A&M UNIVERSITY ----------							
TERMIO	141743	0.37928	0.18440	0.00000	0.87646	0.0006973	53760
MAXSTOR	141743	11162.48888	2935.22177	6960.00000	18248.00000	11.0986151	1582204563
DISKIO	141743	12.58826	13.66378	0.00000	474.00000	0.0516653	1784298
CPUSEC	141743	0.01566	0.00485	0.01013	0.72893	0.0000183	2219
TWOKMINS	141743	0.02510	0.02904	0.00066	1.16374	0.0001098	3558
RESPONSE	141743	0.11193	0.09315	0.02000	6.86000	0.0003522	15865
---------- INST=UNIVERSITY OF CENTRAL FLORIDA ----------							
TERMIO	507238	0.38569	0.18174	0.00000	0.88477	0.0003279	195636
MAXSTOR	507238	11180.56092	2887.94474	6960.00000	18200.00000	5.2109106	5671204892
DISKIO	507238	12.98947	13.84833	0.00000	483.00000	0.0249875	6588752
CPUSEC	507238	0.01551	0.00365	0.01029	0.28687	0.0000066	7867
TWOKMINS	507238	0.02546	0.02835	0.00078	1.19435	0.0000512	12914
RESPONSE	507238	0.10222	0.07683	0.02000	4.51000	0.0001386	51849
---------- INST=SANTA FE COMMUNITY COLLEGE ----------							
TERMIO	99153	0.32983	0.16993	0.00000	0.82227	0.0010766	32703
MAXSTOR	99153	11091.66938	3028.85015	6960.00000	17808.00000	19.1903195	1099772240
DISKIO	99153	10.98780	7.68073	0.00000	396.00000	0.0486639	1089473
CPUSEC	99153	0.01543	0.00341	0.01023	0.08302	0.0000216	1530
TWOKMINS	99153	0.02229	0.01789	0.00066	0.97234	0.0001133	2210
RESPONSE	99153	0.10456	0.07627	0.02000	2.99000	0.0004832	10367

Figure 11B. LIX15 Output for November 1988: Transactions by Institution Group.

VARIABLE	N	MEAN	STANDARD DEVIATION	MINIMUM VALUE	MAXIMUM VALUE	STD ERROR OF MEAN	SUM
INST=SUS EXTENSION LIBRARY							
TERMIO	12342	0.28235	0.18565	0.00000	0.80273	0.004201	3484
MAXSTOR	12342	11820.23963	3494.14661	6960.00000	17624.00000	79.066043	145885386
DISKIO	12342	10.16283	11.20515	6.00000	321.00000	0.253552	125429
CPUSEC	12342	0.01540	0.00360	0.01023	0.06016	0.000081	190
TWOKMINS	12342	0.02214	0.02672	0.00736	0.77878	0.000605	273
RESPONSE	12342	0.09528	0.07695	0.02000	1.36000	0.001741	1175
INST=FLORIDA ATLANTIC UNIVERSITY							
TERMIO	498594	0.39248	0.18412	0.00000	0.88037	0.0003827	195688
MAXSTOR	498594	11065.31488	2880.16656	6960.00000	18200.00000	5.9871348	5517099182
DISKIO	498594	13.18764	14.49129	0.00000	477.00000	0.0301237	6575278
CPUSEC	498594	0.01557	0.00485	0.01013	0.91422	0.0000101	7763
TWOKMINS	498594	0.02584	0.02993	0.00078	1.17092	0.0000622	12883
RESPONSE	498594	0.10370	0.08844	0.02000	7.18000	0.0001838	51704
INST=FLORIDA CENTER FOR LIBRARY AUTOMATION							
TERMIO	40907	0.28085	0.22055	0.00000	0.53516	0.02451	11488
MAXSTOR	40907	9652.14815	3252.35425	6960.00000	16232.00000	361.37269	394840413
DISKIO	40907	7.33333	2.68328	6.00000	15.00000	0.29814	299984
CPUSEC	40907	0.01341	0.00309	0.01029	0.02258	0.00034	548
TWOKMINS	40907	0.01388	0.01065	0.00736	0.04247	0.00118	567
RESPONSE	40907	0.05370	0.04716	0.02000	0.23000	0.00524	2196

Figure 11B. LIX15 Output for November 1988: Transactions by Institution Group (cont.).

VARIABLE	N	MEAN	STANDARD DEVIATION	MINIMUM VALUE	MAXIMUM VALUE	STD ERROR OF MEAN	SUM
---------- INST=FLORIDA INTERNATIONAL UNIVERSITY ----------							
TERMIO	760298	0.40075	0.18507	0.00000	0.88477	0.0003503	304689
MAXSTOR	760298	11318.29769	2949.64466	6960.00000	25328.00000	5.5835349	8605278537
DISKIO	760298	13.64661	15.33969	0.00000	1494.00000	0.0290373	10375490
CPUSEC	760298	0.01590	0.00517	0.01018	0.94221	0.0000098	12088
TWOKMINS	760298	0.02736	0.03341	0.00066	4.10753	0.0000632	20801
RESPONSE	760298	0.10758	0.09473	0.02000	8.62000	0.0001793	81792
---------- INST=FLORIDA STATE UNIVERSITY ----------							
TERMIO	1179775	0.41654	0.18398	0.00000	0.87988	0.0002318	491423
MAXSTOR	1179775	11562.26829	2996.70697	6960.00000	50656.00000	3.7752832	13640874677
DISKIO	1179775	14.72226	28.03290	0.00000	7725.00000	0.0353161	17368954
CPUSEC	1179775	0.01604	0.00645	0.01021	1.32625	0.0000081	18923
TWOKMINS	1179775	0.02975	0.11725	0.00066	52.31238	0.0001477	35098
RESPONSE	1179775	0.10201	0.08985	0.02000	11.08000	0.0001132	120348
---------- INST=UNIVERSITY OF NORTH FLORIDA ----------							
TERMIO	464770	0.39862	0.17715	0.00000	0.88184	0.0004263	185266
MAXSTOR	464770	11366.22172	3009.94954	6960.00000	26896.00000	7.2433497	5282678530
DISKIO	464770	13.57993	15.17369	0.00000	906.00000	0.0365150	6311544
CPUSEC	464770	0.01592	0.00405	0.01021	0.32193	0.0000097	7399
TWOKMINS	464770	0.02726	0.03197	0.00066	2.35543	0.0000769	12669
RESPONSE	464770	0.10899	0.09070	0.02000	11.25000	0.0002183	50655

Figure 11B. LIX15 Output for November 1988: Transactions by Institution Group (cont.).

VARIABLE	N	MEAN	STANDARD DEVIATION	MINIMUM VALUE	MAXIMUM VALUE	STD ERROR OF MEAN	SUM
---------- INST=UNIVERSITY OF SOUTH FLORIDA ----------							
TERMIO	1984281	0.41844	0.18820	0.00000	1.37939	0.0002167	830302
MAXSTOR	1984281	11708.18836	3013.96293	6960.00000	20712.00000	3.4697594	23232334957
DISKIO	1984281	15.13892	18.11746	0.00000	1797.00000	0.0208573	30039871
CPUSEC	1984281	0.01641	0.00615	0.01013	1.54841	0.0000071	32562
TWOKMINS	1984281	0.03113	0.04026	0.00066	4.75556	0.0000463	61770
RESPONSE	1984281	0.12270	0.10121	0.02000	8.98000	0.0001165	252003
---------- INST=UNIVERSITY OF FLORIDA ----------							
TERMIO	3342559	0.41557	0.18815	0.00000	0.88818	0.0001672	1389067
MAXSTOR	3342559	12000.58716	3112.63120	6960.00000	42280.00000	2.7667811	40112676061
DISKIO	3342559	21.24255	140.35919	0.00000	10455.00000	0.1247636	71004476
CPUSEC	3342559	0.01753	0.01996	0.01018	2.33945	0.0000177	58595
TWOKMINS	3342559	0.04996	0.40706	0.00066	33.23778	0.0003618	166994
RESPONSE	3342559	0.11924	0.13491	0.02000	13.49000	0.0001199	398566
---------- INST=UNIVERSITY OF WEST FLORIDA ----------							
TERMIO	599149	0.39741	0.18071	0.00000	0.87305	0.0003713	238107
MAXSTOR	599149	11499.33045	2999.19694	6960.00000	18224.00000	6.1628173	6889812783
DISKIO	599149	14.25136	16.55195	0.00000	462.00000	0.0340113	8538688
CPUSEC	599149	0.01612	0.00444	0.01029	0.69622	0.0000091	9658
TWOKMINS	599149	0.02930	0.03743	0.00066	1.20081	0.0000769	17555
RESPONSE	599149	0.12124	0.10839	0.02000	11.48000	0.0002227	72640

Figure 11B. LIX15 Output for November 1988: Transactions by Institution Group (cont.).

VARIABLE	N	MEAN	STANDARD DEVIATION	MINIMUM VALUE	MAXIMUM VALUE	STD ERROR OF MEAN	SUM
INST=NEWSPAPER INDEX							
TERMIO	16648	0.29709	0.21437	0.00000	0.84375	0.002643	4945
MAXSTOR	16648	10828.04559	2836.95359	6960.00000	16000.00000	34.970858	180265294
DISKIO	16648	12.97234	13.34094	6.00000	117.00000	0.164453	215963
CPUSEC	16648	0.01552	0.00323	0.01029	0.11461	0.000040	258
TWOKMINS	16648	0.02422	0.02321	0.00736	0.28531	0.000286	403
RESPONSE	16648	0.08233	0.08815	0.02000	5.67000	0.001087	1370
INST=UNION AUTHOR/TITLE INDEX							
TERMIO	194957	0.41852	0.22230	0.00000	0.88818	0.0009891	81593
MAXSTOR	194957	12864.12827	3362.02144	6960.00000	22048.00000	14.9582815	2507951850
DISKIO	194957	14.72766	52.59177	6.00000	5352.00000	0.2339909	2871260
CPUSEC	194957	0.01778	0.01508	0.01021	1.18594	0.0000671	3466
TWOKMINS	194957	0.03527	0.14488	0.00736	15.14322	0.0006446	6876
RESPONSE	194957	0.12921	0.12364	0.02000	7.76000	0.0005501	25190
INST=DATABASE SELECTION MENU							
TERMIO	21995	0.40204	0.07811	0.03613	0.54980	0.001675	8842
MAXSTOR	21995	21227.19043	3180.89403	8160.00000	26280.00000	68.221212	466892041
DISKIO	21995	17.01610	3.78071	0.00000	18.00000	0.081086	374269
CPUSEC	21995	0.02846	0.00405	0.01000	0.05266	0.000087	625
TWOKMINS	21995	0.06530	0.01501	0.00066	0.08277	0.000322	1436
RESPONSE	21995	0.15937	0.12760	0.02000	4.18000	0.002737	3505

Figure 11B. LIX15 Output for November 1988: Transactions by Institution Group (cont.).

VARIABLE	N	MEAN	STANDARD DEVIATION	MINIMUM VALUE	MAXIMUM VALUE	STD ERROR OF MEAN	SUM
			FUNCTION=CIRCULATION				
TERMIO	1268423	0.15591576	0.13588201	0.0000000	0.914063	0.00049802	116069
MAXSTOR	1268423	8283.21605502	1922.36629840	4408.0000000	31568.000000	7.04565371	6166357360
DISKIO	1268423	20.26781205	25.11756477	0.0000000	807.000000	0.09205824	15088170
CPUSEC	1268423	0.01713840	0.00663729	0.0084375	0.094193	0.00002433	12758
TWOKMINS	1268423	0.03109094	0.04145760	0.0003224	4.170531	0.00015195	23145
RESPONSE	1268423	0.27503479	0.24010257	0.0200000	15.230000	0.00088000	204746
			FUNCTION=LUIS				
TERMIO	6370171	0.4061325	0.18575724	0.0000000	0.884766	0.00023420	2554870
MAXSTOR	6370171	11627.6675807	3051.85696358	6960.0000000	26976.000000	3.84781029	73146517280
DISKIO	6370171	15.9267351	65.39779844	0.0000000	10452.000000	0.08245417	100190790
CPUSEC	6370171	0.0167953	0.01491776	0.0102344	2.339453	0.00001881	105654
TWOKMINS	6370171	0.0341392	0.18835339	0.0006641	29.688615	0.00023748	214761
RESPONSE	6370171	0.1185577	0.11777748	0.0200000	13.490000	0.00014849	745815
			FUNCTION=TECH SERVICES				
TERMIO	2031840	0.2535243	0.21705096	0.0000000	1.461426	0.00040188	739520
MAXSTOR	2031840	11806.4872470	3089.25854555	5104.0000000	29192.000000	5.71990769	34439051040
DISKIO	2031840	5.5696170	13.87984379	0.0000000	1314.000000	0.02569918	16246350
CPUSEC	2031840	0.0150429	0.01098458	0.0074740	2.040234	0.00002034	43879
TWOKMINS	2031840	0.0137660	0.03152002	0.0003693	3.253111	0.00005836	40155
RESPONSE	2031840	0.1453728	20.03564342	0.0200000	10741.950000	0.03709694	424046

Figure 11C. LIX15 Output for November 1988: Transactions by Function.

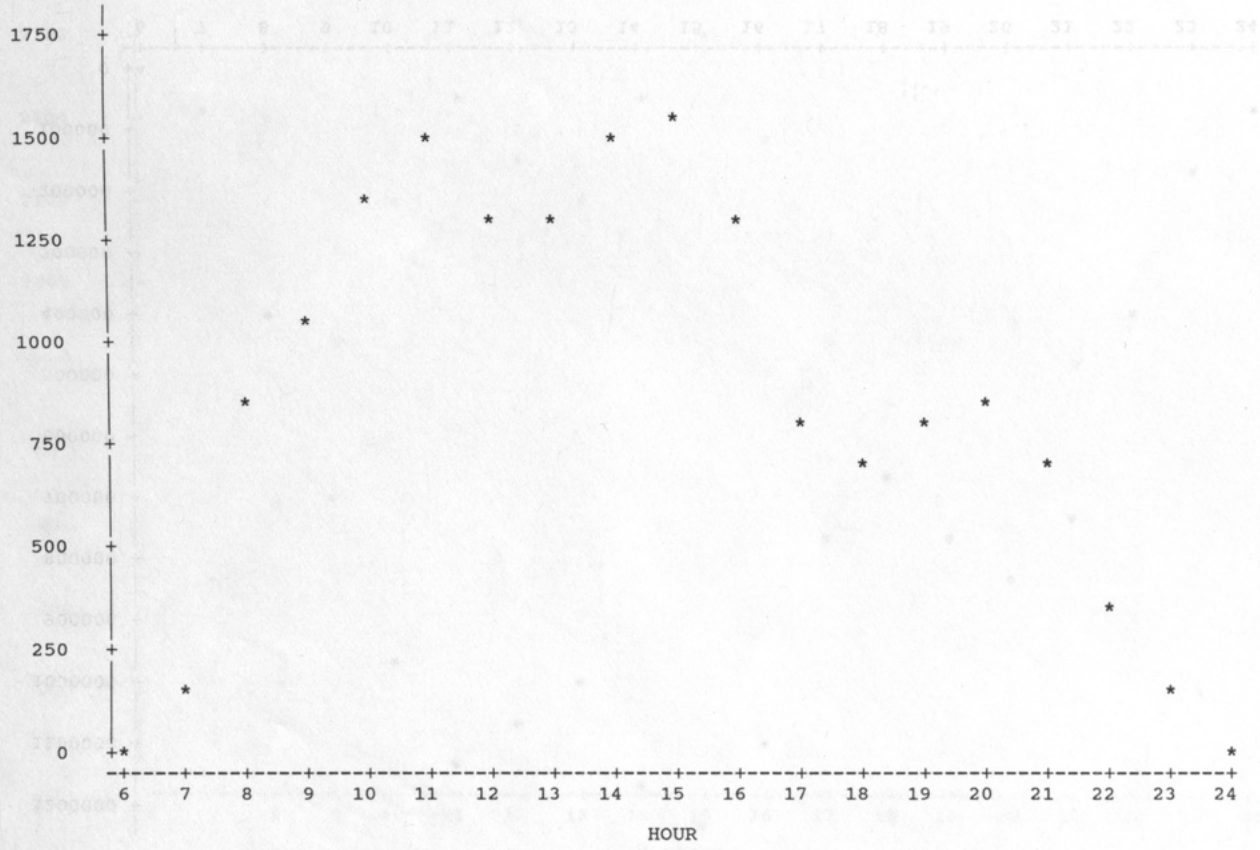


Figure 11D. Total CPU Usage (in Seconds) by Hour of Day.

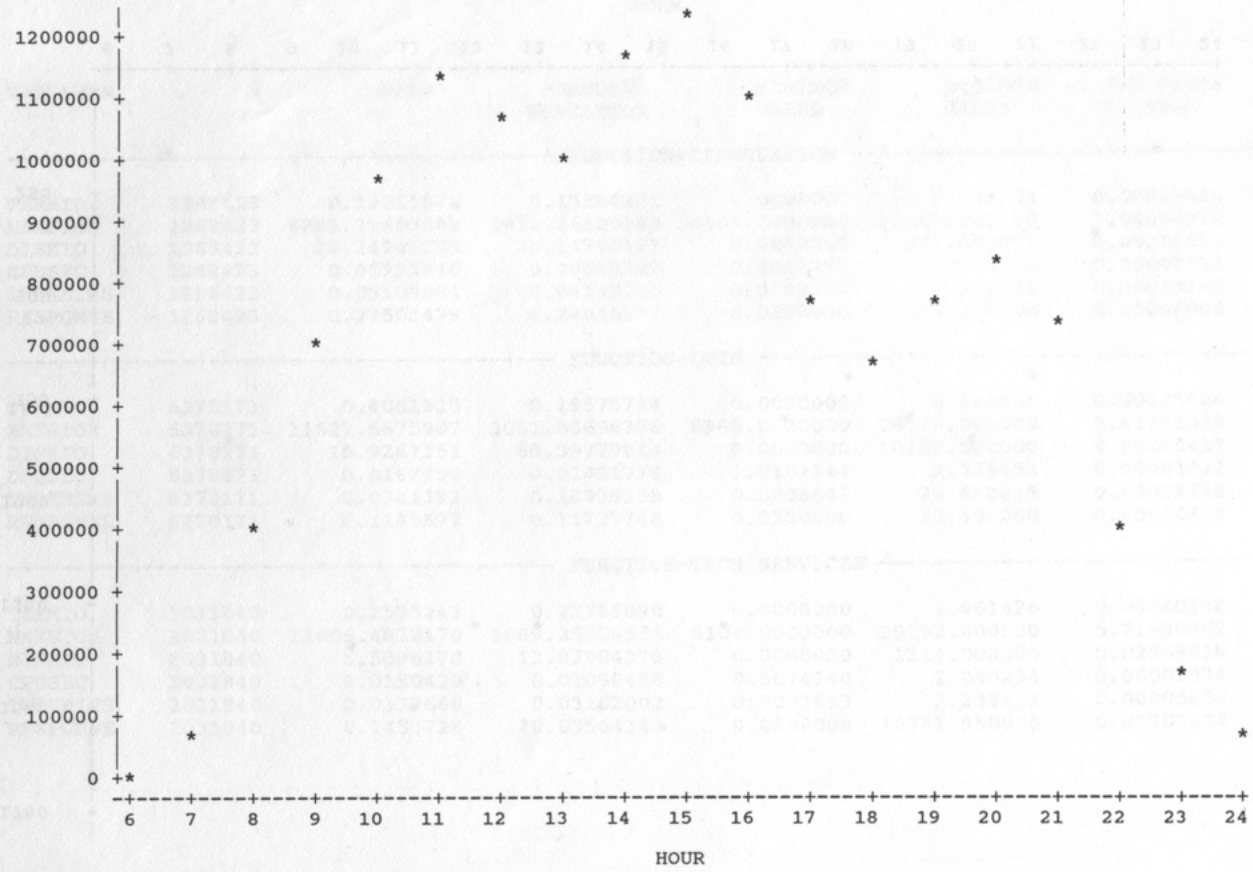


Figure 11E. Total DISK I/O Activity (EXCPs) by Hour of Day.

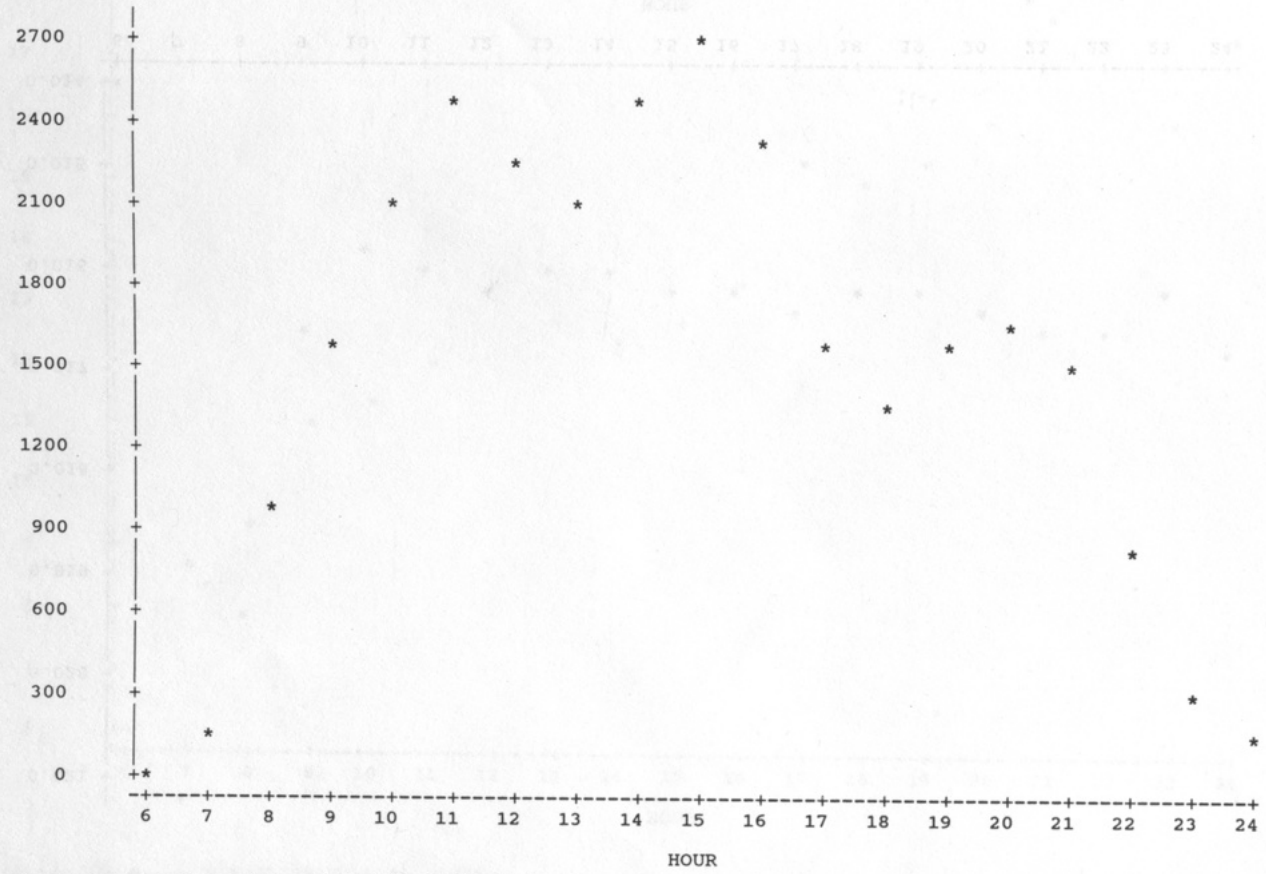


Figure 11F. Total Core Residence (in 2K Minutes) by Hour of Day.

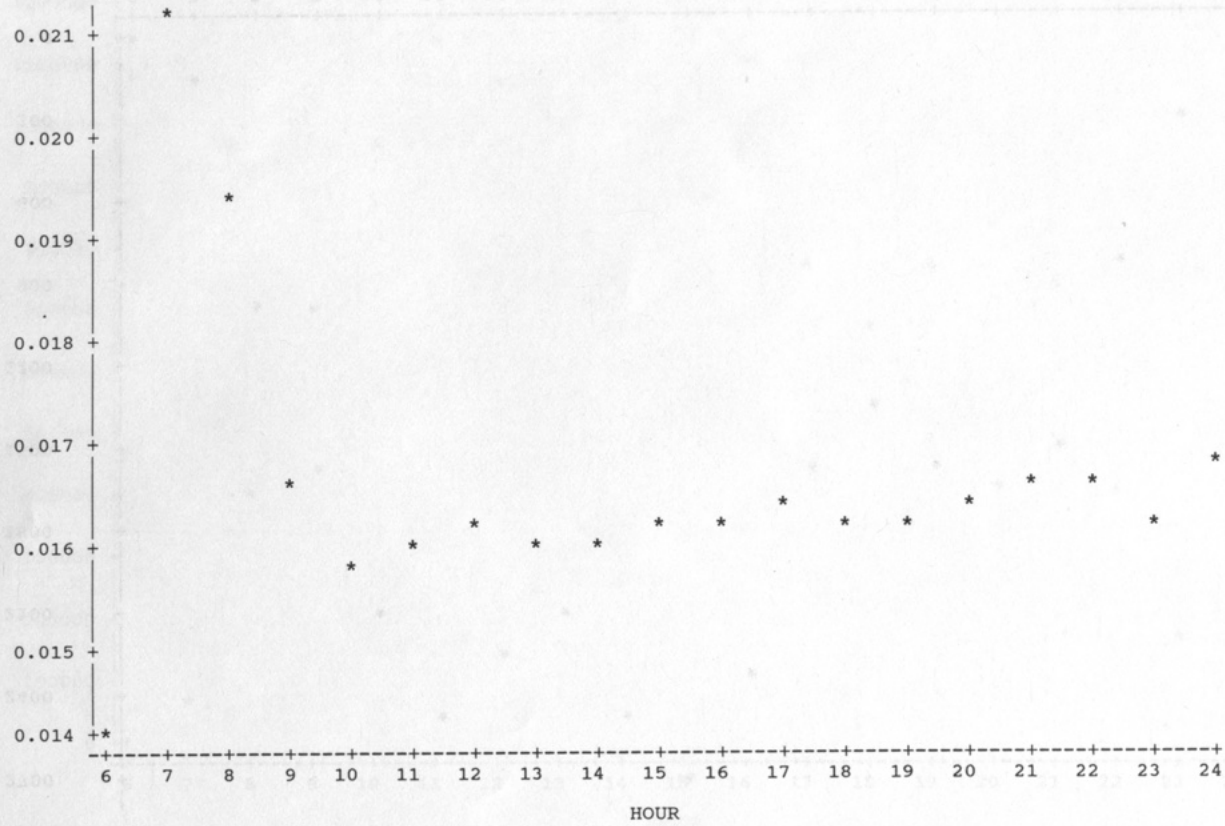


Figure 11G. Average CPU Usage (in Seconds) by Hour of Day.

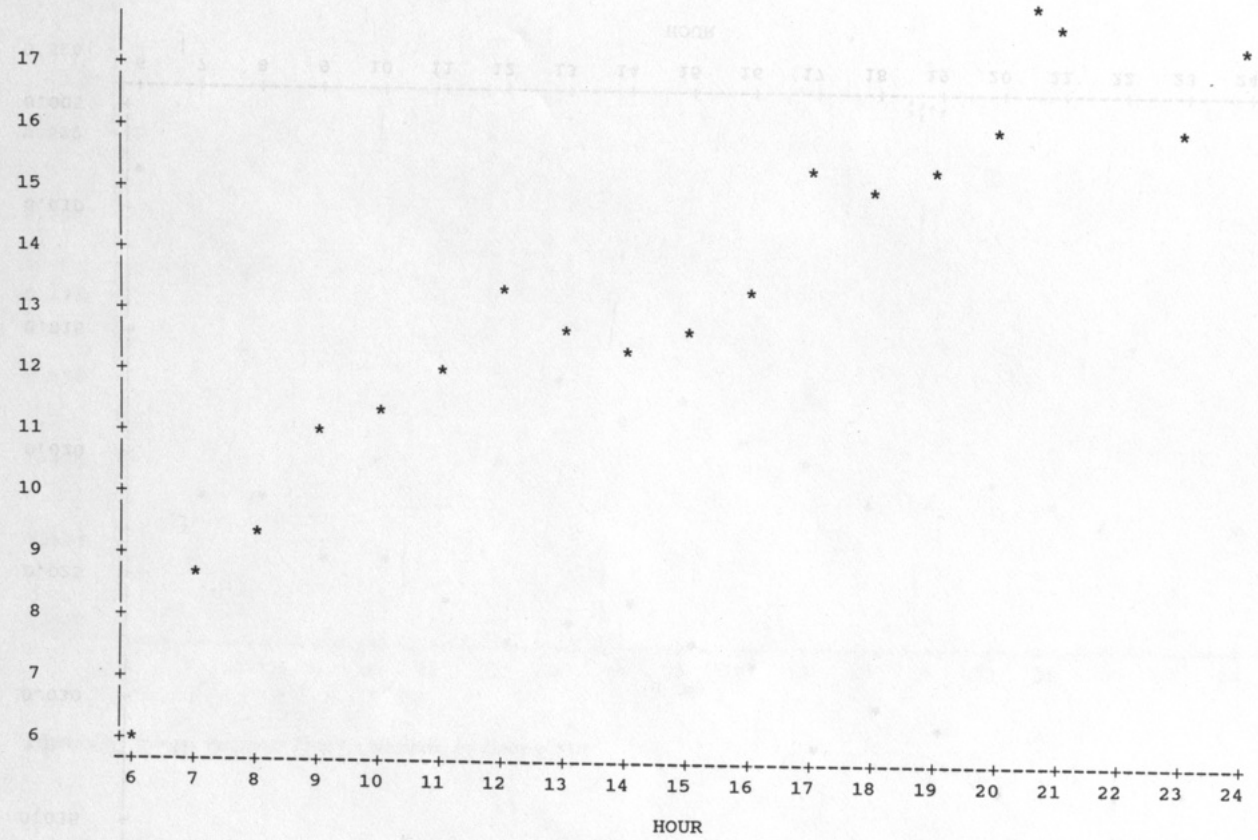


Figure 11H. Average Disk I/O (EXCPs) by Hour of Day.

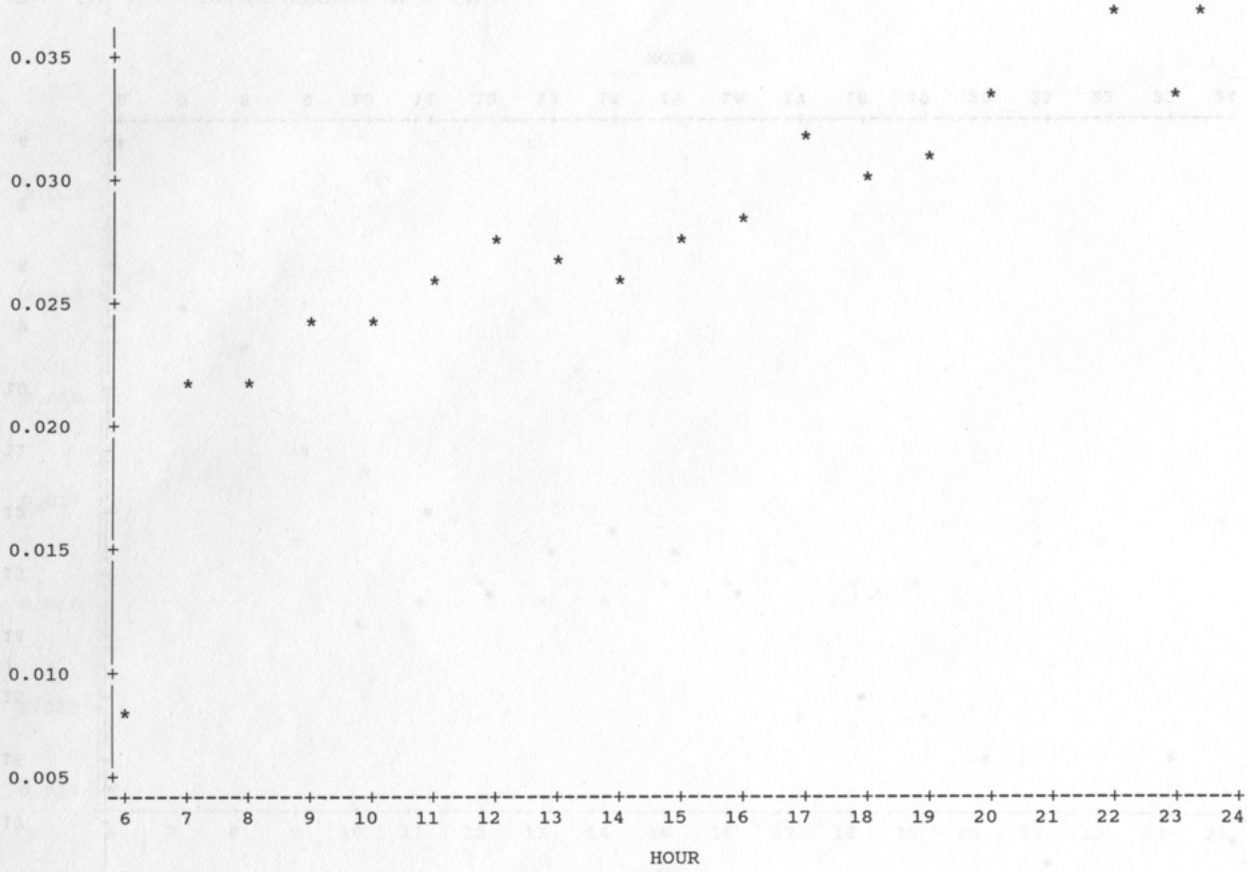


Figure 111. Average Core Residence (in 2K Minutes) by Hour of Day.

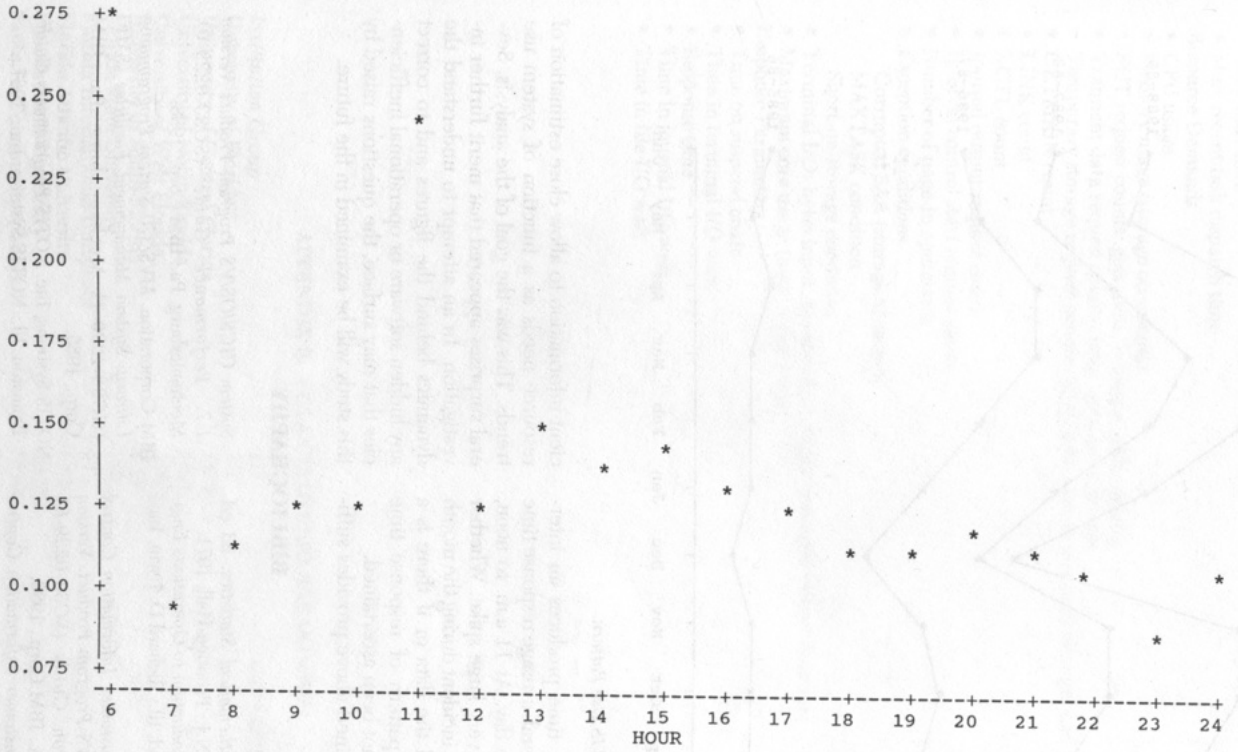


Figure 11J. Average Response Time (in Seconds) by Hour of Day.

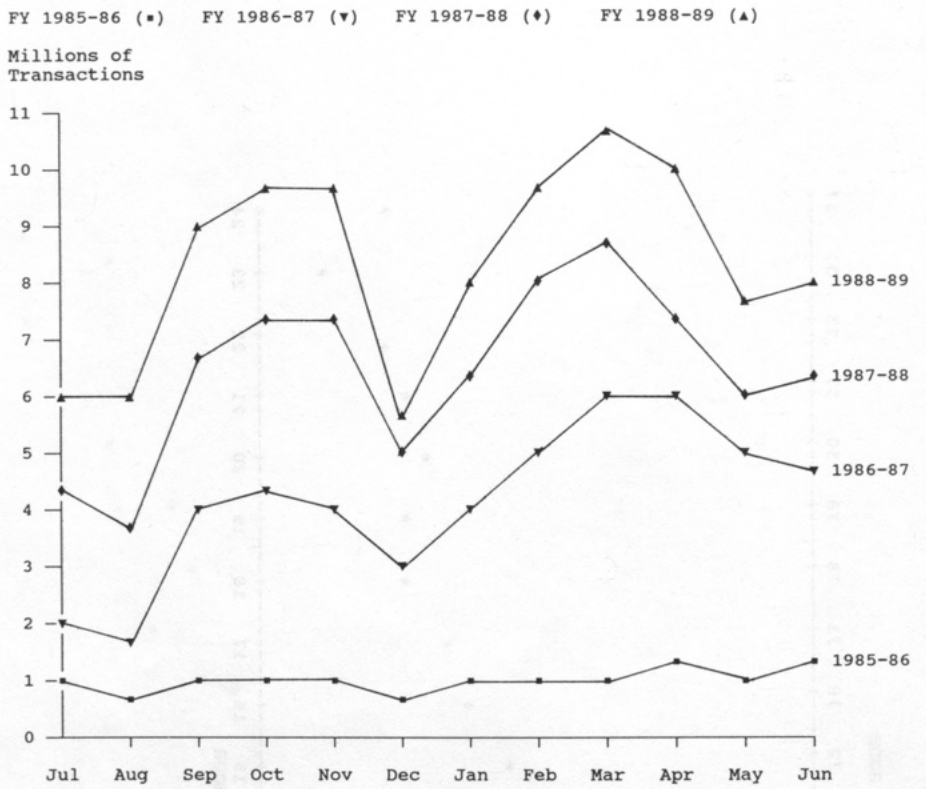


Figure 12. FCLA NOTIS Use Pattern.

Average response time produces an interesting shape. In general, average response time varies little over the day. At 11 a.m. to noon, however, there is a very large spike. Whether this is due to a single incident during the month that heavily skewed the data or if there is a regular day-to-day pattern of response time delays at noon has not been ascertained.

The analysis outlined above provides sufficient information to allow close estimation of resource needs as a function of system use trends. This was the goal of the analysis. Several surprises appeared that merit further investigation. In an attempt to understand the dynamics behind the figures and to correct any hidden software or operational inefficiencies that may surface, the questions raised by this study will be examined in the future.

BIBLIOGRAPHY

Freund, John E. *Mathematical Statistics.* 2d ed. Englewood Cliffs, N.J.: Prentice-Hall, 1971.

Giffen, Walter C. *Introduction to Operations Engineering.* Homewood, Ill.: Richard D. Irwin, Inc., 1971.

IBM Corporation. *Customer Information Control System CICS/OS/VS Program Product. Version 1.7. Customization Guide (SC33-0239-0).* Mechanicsburg, Pa.: IBM Corp., 1986.

IBM Corporation. *Customer Information Control System CICS/OS/VS Program Product Version 1.7. Performance Guide (SC33-0229-0).* Mechanicsburg, Pa.: IBM Corp., 1988.

IBM Corporation. *MVS/370 System Programming Library; System Management Facilities (SMF) (GC28-1030-4).* Mechanicsburg, Pa.: IBM Corp., 1985.

NOTIS Systems, Inc. *NOTIS Programmers Guide.* Evanston, Ill.: NOTIS Systems, Inc., 1987.

APPENDIX A. AVAILABLE TRANSACTION LEVEL CMF PERFORMANCE METRICS

Identification
- Transaction ID
- Task type
- Terminal NETNAME
- Task number
- Operator ID
- Task start time
- Most recent task dispatch time

Resource Demands
- CPU usage
- Memory usage (storage occupancy)
- FCT request counts: gets, puts, browses, adds, deletes
- Transient data request counts: total, gets, puts, purges
- Temporary storage request counts: total, gets, puts to auxiliary storage, and puts to main storage
- GETMAIN count
- LINK count
- XCTL count
- Journal output request count
- Interval control start request count
- Number of page-in operations
- Exception conditions
  - Corrupted SAA (storage violation)
  - MAXTASK condition
  - Short-on-storage condition
- Terminal I/O: bytes input, bytes output, input messages, output messages
- Maximum core usage (high-water mark)

Response Parameters
- Time on suspend chain
- Time in terminal I/O wait
- Response time
- Time in journal I/O wait
- Time in file I/O wait

APPENDIX B. FCLA CIS OPERID BREAKDOWN

Institution Group	*ID*	*CICS OPERID 1st Character*
University of Florida	UF	A
University of Central Florida	UCF	B
University of North Florida	UNF	C
University of West Florida	UWF	D
University of South Florida	USF	E
Florida State University	FSU	F
Florida International University	FIU	G
Florida Atlantic University	FAU	H
Florida A&M University	FAMU	I
Santa Fe Community College	SFCC	J

Processing Unit	*ID*	*CICS OPERID 2d Character*
UF Main	UF	A
UF Health Center	HC	B
UF Law	LL	C
UCF Main	CF	A
UNF Main	NF	A
UWF Main	WF	A
USF Main	SF	A
USF Medical	SM	B
USF Ft. Myers/Edison CC	EC	C
USF Sarasota	SA	D
FSU Main	FS	A
FSU Law	FL	B
FSU CRL	CR	C
FIU Main	FI	A
FAU Main	FA	A
FAMU Main	AM	A
Santa Fe CC	CC	A

Operator	*CICS OPERID 3d Character*
Acquisitions	A
Acquisitions Supervisor	O
Cataloging	B
Cataloging Supervisor	1
Circulation Desk	C
Circulation Supervisor	2
Institution Supervisor	D
Vendor File Maintenance	E
Serials Check-in	F
Library Staff OPAC Access	G
Item File Maintenance	H
Authorities File Maintenance	I
Remote Access	X,Y,Z

Networking and Authority Control: Online Catalog Authority Control in Illinois

Barbara Henigman

As libraries implement online public access catalogs, the need for online authority control is becoming increasingly clear. This need is even more essential as libraries link databases or create multilibrary shared catalogs. In facing the challenge of maintaining authority control for the Illinois statewide online union catalog, ILLINET Online, libraries throughout Illinois have developed a system for shared authority work among several sites throughout the state. This system has merged the policies and work flow of nine individual libraries so authority control can be achieved at a statewide level. The experience has shown that libraries can successfully combine efforts to develop policies and procedures that standardize the sharing and use of data.

Online authority control is probably the most important consideration for libraries implementing online catalogs. Indeed, "online authority control is at the forefront of library computer technology, and many libraries are grappling with new procedures, major projects, and reorganization of work flow brought on by practicing authority control for online, rather than manual, catalogs."[1] A recent survey of thirty ARL libraries showed that a majority of these libraries are using or plan to implement some type of online authority control.[2]

With the capabilities of consistent retrieval made possible by a system of online authority control, many libraries are now investigating the creation of shared catalogs. This is done in order to take full advantage of resource sharing and cooperative collection development. Along with the concept of a shared database comes an even greater need for online authority control in order to keep the level acceptable for shared use. The concept of a shared database also poses the challenge of allowing more than one participant in the network to take on the responsibilities of authority control.

The challenge of shared authority control has been addressed by libraries in Illinois, beginning in March of 1989 when the ILLINET Online database (IO) was completed and designated to function as a statewide online union catalog for Illinois. IO's more than five million unique bibliographic records represent the entire OCLC cataloging history of all Illinois OCLC member libraries. The database is currently updated weekly, providing continuous additions of all new materials that are cataloged statewide. Based on software acquired in 1980 from the Western Library Network (WLN), this particular system was chosen for use in Illinois because it was designed to function within a network environment and it allowed for full authority control including a syndetic subject structure.[3] From the beginning, librarians realized that a system of authority control would be vital to the successful implementation of a statewide online union catalog in Illinois.

Barbara Henigman is the Assistant Automated Systems Maintenance Librarian and Assistant Professor of Library Administration at the University of Illinois at Urbana-Champaign.

NEED FOR AUTHORITY CONTROL

The anticipated size, scope, and growth rate of the database mandate shared online authority control. The scope of participants in the network reflects many types of libraries including public, academic, special, and school libraries. At present there are approximately 800 individual libraries whose holdings are represented in IO. Many of these participants use IO for resource sharing and continue to maintain in-house catalogs. However, others rely on IO as the only means of access to their collections. Combining resources of many different types of libraries has created a need to reconcile vastly different interpretations of authority control into one shared concept.

Since IO's authority file is searchable by the user and is designed to provide the user with a direct link to bibliographic data, a system of authority control is essential to maintain a high level of retrievability for users. At present seven of the major academic libraries in Illinois have closed or are in the process of closing their card catalogs, relying solely on IO for access to their collections. While the bibliographic data in the database can be limited to the holdings of their individual libraries, the online authority file is shared statewide. The sheer size of this shared authority file makes authority control an absolute necessity if these libraries are to provide effective user services. The many other libraries that use IO for interlibrary loan activity also rely on complete and accurate retrieval if they are to receive maximum resource sharing benefits.

The maintenance of a shared authority file for the ILLINET Online database required the development of authority control procedures that could be implemented statewide to achieve consistency in work flow, staff training, and cataloging policy. The task proved to be complex, since it involved coordinating both the cataloging policies and authority control procedures of libraries that otherwise had no administrative relationship to one another. The procedures presently in place were carefully designed to foster a precise understanding of the structure of the database, a strong commitment to communication and policy development, and an ongoing commitment to training and current documentation.

AN AUTHORITY FILE-BASED RETRIEVAL SYSTEM

The ILLINET Online database was designed as a union catalog for the purpose of providing users access to unique bibliographical entities rather than displaying variations generated when an item is cataloged by several libraries. Only one catalog record exists for each title cataloged. The database uses the OCLC control number as a record identifier to identify duplicate records as they enter the database. The IO record usually represents the cataloging done by the first Illinois library to catalog the record through OCLC. Subsequent variations of the catalog record are rejected.

The ILLINET Online database was created from OCLC tapes containing the cataloging done by Illinois OCLC member libraries from 1975–1989. Created in two phases, the first phase involved the creation of a pilot database that would demonstrate the feasibility of a statewide online union catalog. In 1984, a database of approximately 750,000 bibliographic records was loaded by combining the OCLC records from 1975–1984 of the University of Illinois at Urbana-Champaign and the River Bend Library System (a consortium of twenty-eight public, school, and small academic libraries in northwestern Illinois). The second phase began in 1988 and added the cataloging of the remaining OCLC member libraries in Illinois, creating a database of over four million bibliographic records. IO is updated weekly by running OCLC tapes of the cumulated cataloging from member libraries against the existing database.

IO stores the complete MARC record and presently accepts records in all eight MARC formats—books, serials, scores, sound recordings, audio-visual media, maps, manuscripts, and computer files. Once a catalog record enters the IO database, it undergoes some transformation to allow it to be stored within the ILLINET Online design. The database divides the information from the MARC record into three files: the Bibliographic File, Authority File, and Holdings File. The breakdown of these files is shown in figure 1. As is typical of any algorithmic-based device that must work with consistent and strictly defined values, data from the catalog record are sorted strictly on the basis of MARC tags, not on the content of the fields themselves.

The bibliographic file contains those coded areas of the catalog record that are traditionally considered to represent descriptive information about an individual item. The holdings file contains the name of the library holding an individual item. Since most participating libraries have their own circulation systems, local call numbers and other information are not present in IO. At present, users have a direct link to the circulation information of thirty-eight libraries that use the same circulation system, and a future goal is to provide a direct link to other circulation systems in Illinois. The authority file contains headings or access points used in retrieving the descriptive information. This structure allows a heading or access point to be stored only once and still be accessed by any number of unique bibliographic records that may be related to that heading. This division of information remains invisible to the user since each individual bibliographic record is displayed with its associated headings as a single entity through commands that are given by the user.

As bibliographic records enter the database, headings are compared to those already in the authority file. If a heading on the incoming bibliographic record already exists in the IO authority file, the program simply creates a "bridge" between the descriptive information (bibliographic file) and the already existing authority record (authority file). If a heading does not match any existing authority record, the program will create a new authority record and build the "bridge." When corrections are made to a heading in the IO authority file, the editing program automatically creates a new "bridge" for all bibliographic records that were previously linked to an incorrect record. This global bibliographic change capacity makes it possible to achieve consistency in the bibliographic file by maintaining the authority file.

Since the database design for the IO authority file was written by WLN before the LC MARC Authorities format was developed, the IO authority file structure contains some unique features that keep it from being compatible with the present LC structure. As shown in figure 1, the IO authority file consists of three subfiles, each containing a specific type of heading. Just as the division of bibliographic data vs. authority data is decided strictly on the basis of MARC tagging, the categorization of the three types of headings in the authority file subfiles is selected based on tagging from the original University of Illinois at Urbana-Champaign (UIUC) bibliographic record. The breakdown of these subfiles is also shown in figure 1.

Because the headings are generated from bibliographic input, the most common types of errors are variations in the form of names. Since the database was built from the entire backlog of Illinois OCLC cataloging, pre-*AACR2* forms of entry exist alongside the current *AACR2* form. This system of creating an authority file has the advantage of linking headings in the authority file with the bibliographic information, allowing authority control in both files. On the other hand, generating authority file headings directly from bibliographic records creates the need for a manually generated correction system, since variations in bibliographic cataloging may produce variant forms of a heading. Also, all cross-references must be keyed in manually.

AUTHORITY WORK

Within a single library, authority work for a database of this design could be developed fairly easily. A manual authority control system for a card-based environment would be transferred to an online environment while making use of the same staff resources and many of the same procedures. Likewise, policy could be molded to fit the online authority file. Since the UIUC began as the first authority control site for the IO network, procedures were developed to provide a prototype for statewide expansion. "Procedures involve eliminating unauthorized authority entries by merging them into one form or deleting them from the database. Name authority headings are verified on OCLC's Library of Congress Name authority file (LCNAF) for *AACR2* form."[4] LC subjects are verified using the latest edition of the *Library of Congress Subject Headings.*

Most work is generated for the UIUC authority control unit through three channels. The highest priority is given to authority work that is generated by the original and copy catalogers. Headings verified during the cataloging process are searched in the IO database. Any variant forms of a heading are reconciled with the established form and cross-references are added. Adding cross-refer-

ences carries a high priority since any subsequent incoming headings that may match an existing cross-reference are automatically converted to the established form. High priority is also given to correction requests from library personnel other than catalogers. All areas of the library are encouraged to submit problems found in both authority records and bibliographic records, provided that the request is substantiated by a copy of the title page, an LC authority record, or other appropriate verification. These requests represent approximately one third of all work done in the UIUC authority control unit. A third priority is assigned to the LC subject heading changes that are published quarterly in the *Cataloging Service Bulletin*. These subject changes are last in the priority order because they are more time-consuming than the others. Since each combination of subdivisions creates its own heading in the IO authority file, to change such a heading as Music, popu-

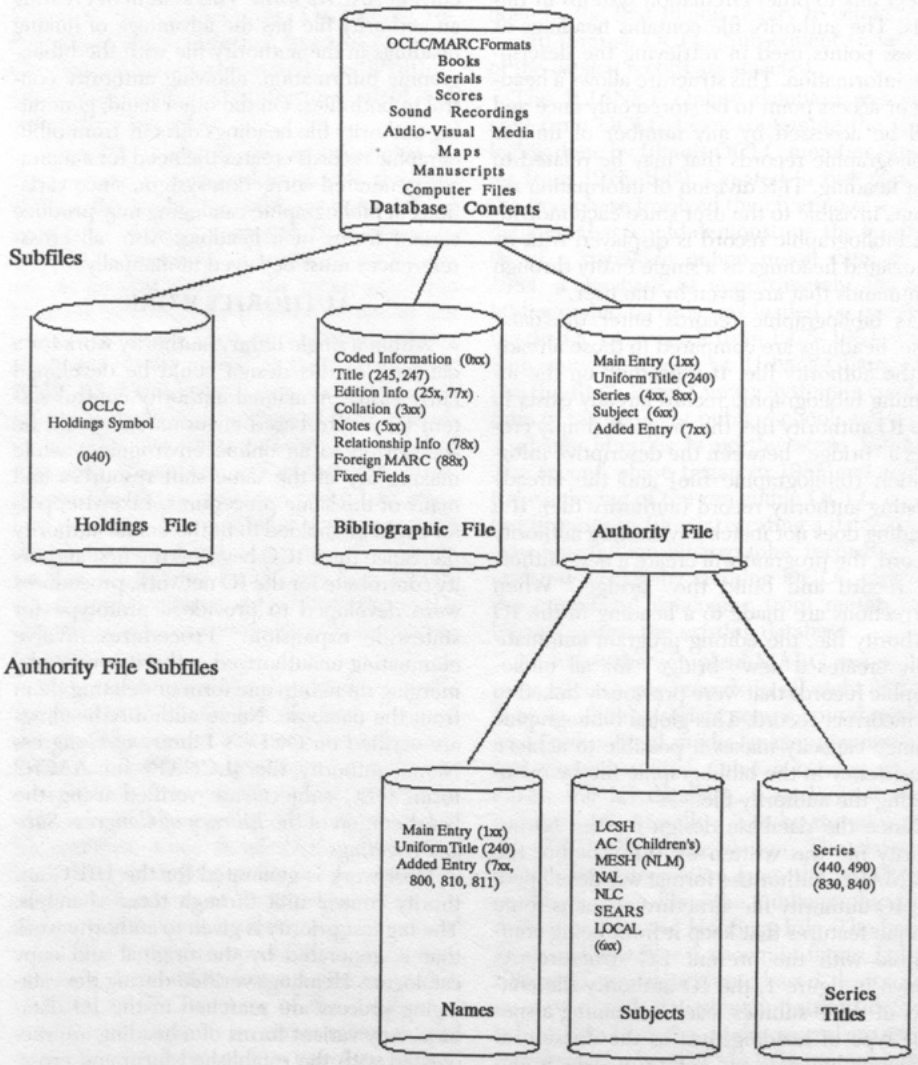


Figure 1. IO Database File Structure.

lar (songs, etc.) to Popular music can be quite time-consuming. The major goal in expanding the authority work procedures into the statewide network was to take several geographically dispersed libraries with their own internal authority control policies and help them work together to provide consistent authority control statewide. Southern Illinois University at Carbondale, Illinois State University, and Northern Illinois University all volunteered to be the first authority control sites. These libraries were logical candidates since all had established in-house authority control procedures and were implementing card catalog closure. The first area of emphasis was placed on training and documentation. Once these sites were fully trained, emphasis was placed on communication and mechanisms for policy-making.

TRAINING

Before personnel could be trained, the UIUC maintenance staff and library system programmers designed an internal system of routing and maintaining data. Records and communications needed to be transferable between sites so that conflicts in form of headings could be discussed and resolved. As a single authority control site, the UIUC system consisted of twelve separate computer work spaces called institutions. The configuration of the institutions is shown in figure 2. The institutions are arranged in a hierarchical configuration to allow the work of a beginning staff member to be reviewed by a senior staff member. Ultimately all edited records are routed to a batch processing program (Verified Record Output or VRO) which runs nightly. "The process of routing records from one 'institution' may access its own group of three working subfiles and each must sign on to a particular 'institution' with a unique password to ensure the protection of the database from unauthorized users."[5] Since the configuration worked well at UIUC, each authority control site was given this internal design. A routing system among authority control sites was implemented so an experienced site could effectively review the work of a beginning site. These are the institutions marked BARBOLD (a locally meaningful acronym) in figure 2. During an initial training period, new authority control sites could work in the institutions marked "Novice" and have their work sent directly to a reviewer at a veteran site who would then send them on for processing.

Once the mechanism for sharing data and resolving conflicts among authority control sites was in place, the training of personnel began. Each designated authority control site sent two staff members—one professional cataloger and one support staff person—to a three-day workshop held at the University of Illinois at Urbana-Champaign. The workshop was designed and conducted by the assistant automated systems maintenance librarian, supervisor of the UIUC authority control staff, and UIUC's principal cataloger. Although intensive, it was designed only as an introduction to the mechanics of the system and basic authority control work flow. It was assumed that all participants had a complete knowledge of cataloging practice and rules. Trainees were given a detailed lecture on database design and command structure. Then each pair of trainees spent a block of time each day actually processing various types of common problems. Training concentrated on resolving conflicts with personal and corporate names, series titles, and subject headings.

Emphasis was also placed on helping each pair of trainees develop a work flow that would reroute present authority work procedures at each library through IO's editing system, creating a parallel to the work flow procedures designed at UIUC. Authority control sites were first encouraged to have original and copy cataloging staffs send their authority work to a central unit for processing against the online authority file. Once a routine was established to handle authority work generated by the catalogers, the work flow system was expanded to allow library personnel from areas other than technical services to submit error reports directly to the corrections unit. Most authority control sites have developed in-house training for their library personnel to facilitate this work flow. The ultimate goal for the statewide network is to provide a mechanism that would allow any IO user to submit an error report to a designated site. Due to the library-wide and statewide concept of error reporting, the concept of a single corrections unit within a technical services department was chosen, rather than expecting each cataloger to make corrections while cataloging.

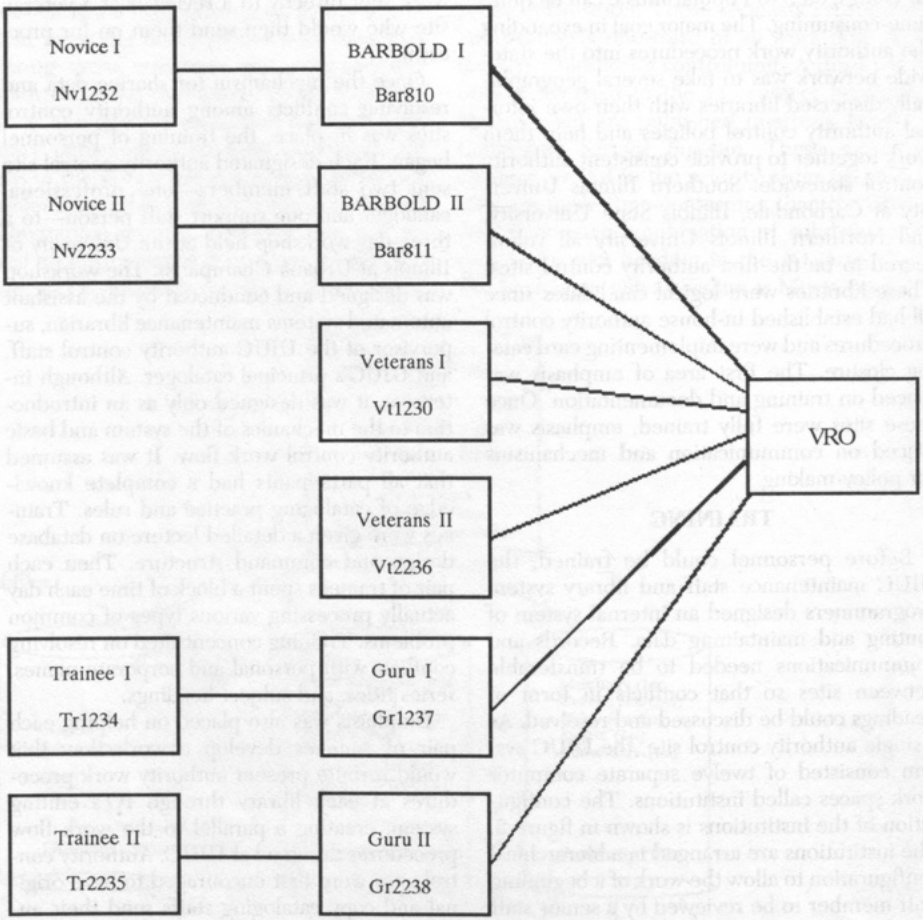


Figure 2. Institution Setups.

DOCUMENTATION

Most of the authority work done by the authority control sites consists of resolving conflicts in form of heading and adding appropriate cross-references. Documentation was designed not as a step-by-step manual of correction procedures, but rather as a reference guide that outlines typical problems and how to solve them given the unique design of the IO authority file. Much of the information is the cumulation of experience gained by the UIUC staff. The manual itself contains short narrative descriptions of problems followed by examples of how the conflicts are resolved. Through examples, the manual describes how to merge the variant forms into one correct form documented by LC and how to add appropriate cross-references. Other documentation includes reference materials such as searching guides, guides to error messages, etc. Once authority control sites are fully trained, they are encouraged to develop in-house documentation that will enhance the ability of their staff and help with internal training programs.

COMMUNICATION AND POLICY MAKING

Once each authority control site was fully trained, ongoing communication among sites was essential to insure that established practices were maintained and that documen-

tation was kept up to date. In order to facilitate this currency, the Illinois Library Computer Systems Organization (ILCSO) Policy Council (the network's governing council) appointed the ILCSO Maintenance Committee to monitor authority control as well as related policy issues throughout the statewide network. The ILCSO Maintenance Committee then appointed the Bibliographic/Authority file Maintenance Sites Subcommittee. The subcommittee meets bimonthly to discuss work flow issues, training progress at individual sites, and to revise documentation. Meetings also serve as a forum to resolve disagreements over how a heading has been established or to work out differences in interpretation of *AACR2* and LC rules. The subcommittee is also responsible for training and monitoring new sites. The need for policy decisions is then brought to the attention of its parent committee.

From the beginning, librarians at authority control sites committed themselves to molding their library cataloging policy to fit into the concept of contributing to a shared statewide union catalog. In order to preserve the integrity of each library's right to participate in the statewide database, concise documentation was written that would clearly define the role of the authority control site. A philosophy of "the more information, the richer the record" has been adopted, making it acceptable to add extra subject headings and added entries and retain locally derived headings as long as they conform to current standards. The primary purpose of the authority control sites is to update form of entry to conform to current *AACR2* guidelines and LC rule interpretations. "What gets changed in IO," a document written by the Bibliographic/Authority File Maintenance Sites Subcommittee, gives specific documentation regarding the treatment of name headings, name-uniform title headings, subject headings, series title headings, and cross-references. As a rule, descriptive elements other than headings are not changed unless a request to do so is accompanied by sufficient documentation to substantiate the request. Minutes from the subcommittee meetings are also used as policy documents when the need arises to modify or add to current procedures. The minutes are indexed and distributed after final approval by the ILCSO Maintenance Committee.

CONCLUSION

After one year of implementation, this method of sharing the needs of authority control is on its way to becoming a successful project. Librarians in Illinois have successfully coordinated their resources toward maintaining a statewide online union catalog. At the same time, they have been able to keep a local focus on the specific needs of their library's users. DePaul University, Northeastern Illinois University, Eastern Illinois University, Western Illinois University, and Governor's State University have all become authority control sites during 1990, bringing the total number of sites to nine. The intent is to allow any library that wishes to make the commitment to statewide authority control to become an authority control site. The ILCSO Maintenance Committee has approved a set of criteria for becoming a new site. Potential authority control sites must be willing to commit twenty hours of staff time per week under the direction of a professional cataloger to correcting the database, be willing to undergo the established training procedure, be willing to participate on the Bibliographic/Authority File Maintenance Sites Subcommittee after training, and agree to work with any special equipment needed to allow staff terminals to be capable of editing the database.

As long as the sites follow the criteria listed above, fully trained sites have the flexibility to develop an internal mechanism that will be compatible with their individual library's work flow. Sites are encouraged to emulate the work flow patterns developed at UIUC, and all current participating sites have developed variations on these patterns. Sites have been given the flexibility of rearranging the priority order for correction requests so that they may still have the opportunity to best serve the specific needs of their users. Sites also are able to develop staffing policies and procedures that will work within their present technical services structures.

A mechanism for error reporting from nonauthority control sites is currently being developed. The Bibliographic/Authority File Maintenance Sites Subcommittee has designed and distributed an error report form to all IO participants and is presently reviewing the response to determine how corrections requests should be distributed among the au-

thority control sites. Several other projects that will relieve some of the manual requirements of the system have also begun. Of primary importance is the partial redesign of the IO authority file structure that will allow the IO authority file to be more compatible with the LC Name authority file (LCNAF). Provision will be made to store the LCNAF control number so LC tapes can be automatically run against the IO authority file. Other future projects involve the use of microcomputer programs to allow automatic updating of routine changes. One example is a program to process changes to the LC subject headings lists that may require minor changes to many subject headings in the database.

Although shared authority control in Illinois is still in its infancy, the success of the initial implementation has shown that several unrelated libraries can cooperate to develop consistent policies and procedures in order to standardize data for sharing. The experience has also shown that Online Authority Control can be extended successfully beyond the boundaries of one library. It is agreed, for the most part, that online catalogs need at least some level of online authority control. "Today there appears to be few adherents of the view that online catalogs need no authority control because truncation, keyword, and Boolean searching will bring together all variations of any name."[6] As libraries continue to build and share online databases, the idea of sharing the responsibilities of authority control may be one solution to the problems presented by online authority control.

REFERENCES

1. Stefanie A. Wittenbach, "Automated Authority Control in ARL Libraries," *Introduction to Automated Authority Control in ARL Libraries, SPEC Flyer 156* (Washington, D.C.: Association of Research Libraries, Office of Management Services, Systems and Procedures Exchange Center, July/Aug. 1989).
2. Ibid.
3. Michael Gorman, "The Online Catalogue at the University of Illinois at Urbana-Champaign: A History and Overview," *Information Technology and Libraries* 4, no.4:306–11 (Dec. 1985).
4. Sharon E. Clark and Winnie Chan, "Maintenance of an Online Catalogue,"*Information Technology and Libraries* 4, no.4:324–38 (Dec. 1985).
5. Ibid., p.332.
6. Jean Dickson and Patricia Zadner, "Authority Control in the Online Environment," *Cataloging & Classification Quarterly* 9, no.3:57–71 (1989). ■■

Communications

Detection and Analysis of Editing Activity in an Online Bibliographic System

Janet Gebbie Edgerton and Raymond G. Taylor

This report is the second of a two-part series on the effectiveness of online editing in an OPAC system. The first part provided a statistical description of the utilization of editing commands; the second part deals only with unsuccessful sessions. In the system studied, most attempts at editing conclude with one or many updates. However, about three percent of the time, no update occurs. The study reported here explains how such failures may be detected and illustrates the circumstances that lead to interrupted or aborted editing sessions.

Most research dealing with transaction logs focuses on how the public uses OPACs. However, this study examines how effectively and efficiently the technical staff in the library uses the editing functions of the OPAC system.

In phase one of the study, the authors presented an overview of editing commands available in North Carolina State University's Bibliographic Information System (BIS) and determined what proportion of the editing sessions were interrupted or aborted (i.e., ended in either a SAVE or CANCEL). Successful sessions normally result in an UPDATE.

In the second phase, as reported here, the authors examined many editing sessions to try to determine why some were interrupted or aborted. Hundreds of thousands of transaction logs were sorted by terminal. Then, working from the sorted file, logs were scanned for SAVEs and CANCELs. When such interruptions or aborted sessions were identified, all of the previous transactions associated with that session were printed. This provided a "trace" of activities that led to the unsuccessful conclusion of the session.

Janet Edgerton is a librarian in the Department of Monographic Cataloging at North Carolina State University. Raymond Taylor is a professor in the Department of Educational Administration and is a member of the program faculty in Operations Research at North Carolina State University, Raleigh.

INTERPRETATION

The mechanical portion of the study was followed by an intensive professional examination of the traces for the purpose of discovering, if possible, the causes of the problems. A graduate library student, together with the first author, studied each trace and, in most cases, interviewed the staff member who generated the trace. For some editing sessions, staff supervisors were also asked to interpret the trace. When possible, the sessions were replicated so that the computer's response to the editing commands used in the unsuccessful session could be observed. Generally, only the most recent logs were used; thus, the editing staff was more likely to remember specific traces or patterns of editing activity.

In some cases, it was impossible to tell what editing had been done to a record. Since only the commands that the editor enters in the top two lines of the screen appear in the command traces, any editing done within the body of the record would appear on the transaction log simply as an ENTER. This would reflect the striking of the ENTER key after the editor inserted or changed some text or data on lines other than the top two.

The researchers had no preconceived notions about the causes of editing failures, but they knew from the phase-one study that the percentage of unsuccessful edits was low—less than three percent. In fact, some of those edits were actually "false drops" that occurred when the analysis program incorrectly perceived certain commands as CANCELs or SAVEs. For example, when an editor neglected to enter a search command, and the first word of the search proper began with "sa" or "ca," then the program registered these as SAVEs and CANCELs.

What emerged was not so much a picture of the reasons for specific editing failures but rather an overview of the patterns of work flows in the different departments. The three departments that edit BIS records are as follows: circulation (two terminals), monographic cataloging (five terminals), and serials cataloging (two terminals).

Circulation staff tended not to use the SAVE command, which puts the record in question into a SAVE file, thus rendering it inaccessible to additional editing. Instead, circulation staff tended to CANCEL when they encountered a problem. They would then report this problem to the supervisor. She would investigate, decide on a course of action, call up the record again, edit it, and UPDATE it. According to the supervisor, the staff had encountered previous problems with the SAVE file and had gradually learned to avoid its use. Also, a large and growing stack of "problem books" served as a visual reminder of the importance of resolving problems as expeditiously as possible.

Circulation sessions were usually longer than sessions in either monographic or serials cataloging. Nearly all circulation editing reflected an ongoing project of the department to enter bar code numbers into the holdings half of the BIS record. This is generally a straightforward procedure: staff members usually enter either a call number or the unique record identification number (the OCLC number); staff then wand in a bar code number using a bar code reader; they physically place a bar code inside the book if it does not already have one.

Occasionally a call number search calls up no records, multiple records, or a record that does not appear to match the book in hand. In any of these situations, someone from the circulation department might try a search by title or author. The traces yielded several instances of CANCELS when no UPDATE had been attempted. In consulting with the circulation supervisor, the researchers theorized that the editor had probably read a bar code number into the record and then realized that there was some problem (e.g., the wrong record had been called up and the bar code had been entered before the editor noticed that the record did not match the book or card in hand). The editor would issue a CANCEL command to delete the bar code number and restore the record to its original unedited condition.

The present study also unearthed a few interesting points, all of which turned out to be idiosyncratic to the system. For example, on two separate occasions a MOVE command was noted. This command was presumably unknown to the staff editors. Upon questioning, the editors could not recall ever using the command. Finally, it was discovered that the MOVE command was inserted into the trace by the system in response to another command: EXPAND. Although such illustrations of idiosyncratic behavior have no generic utility to readers of this article, they help make the point that a detailed study, such as the one undertaken here, teaches the researcher a great deal about the nuances and potential technical difficulties of the system in use.

Most of the records that had been canceled had been updated within ten to twelve days of the cancellation. However, an unknown number of these records may have been updated even earlier, because BIS records only the latest UPDATE and drops all information pertaining to when previous updates were undertaken.

The researchers were disappointed to discover that it is nearly impossible to detect what error(s) caused the record to be saved or canceled by the operator. Replicating searches is not always a sure way to discover what problems were encountered, because those problems are removed once the record is subsequently corrected and updated.

The editing sessions of the monographic cataloging staff tended to be shorter than for circulation. However, there were a few examples of extended sessions during the period when the department was engaged in a special project. This project involved searching and correcting all headings for authors who were represented in an anthology used in freshman English classes. The editor working on this project was new to BIS editing and occasionally made use of the SAVE command when he had a question that needed supervisory review. Monographic cataloging staff tended to use the stacking feature (multiple, sequential commands on a single line) more often than either the circulation or serials staff.

Serials cataloging had the fewest entries between SAVEs and CANCELS. Serials editors used the SAVE command much more

frequently than did editors in either of the other two departments. In many cases, a single record would be edited, saved, called out of save, edited again, and then resaved. Holdings records for serials are much more complex than those for monographs. Staff tended to work on portions of the holdings record, save their changes for supervisory review, then work some more on the same record. Serials staff did not tend to stack commands, but editors used a convenient shortcut method for adding additional volumes to a holdings record. Instead of "ADD VOL at 17" followed by an ENTER command and then "ADD VOL at I8", ENTER, etc., staff members added multiple lines with a single command: ADD VOL at 17 I8 19 20 21 22 ENTER. Serials staff also used a technique whereby an editor indicates that volumes should be added at lines 20.1 20.2 20.3 20.4, etc. An ENTER or REFORMAT command causes the lines to appear in the correct order and numbered with whole numbers.

Due to the complex nature of serial holdings, stacking of commands for changes to holdings for serials could quickly get so complicated so as not to be worth the potential time savings. In other instances, however, serials staff might benefit from stacking. For example, one session consisted of a long string of commands to delete records from the database. Each time the DELETE RECORD command is issued, the system asks for confirmation. By stacking the response to the confirming question, the serials editor could have reduced the time required to delete the records.

Serials staff reported that some error messages can be bypassed which is yet another idiosyncratic, but important, lesson from this study. These error messages disappear by saving the records and then calling them back. This trick may account for some of the SAVE activities observed.

Curiously, in several cases, fields were added or holdings data were edited, and instead of entering either SAVE or UPDATE, the operator would CANCEL. Given the times of the transactions, the researchers have speculated that the operator may have wanted to end the session at break time or at the end of the day. Since the amount of work completed was simply not worth saving, the operator probably would take it up again after the break or on another day.

SUMMARY

This two-phase study, supported in part by the Council on Library Resources, investigated the utilization of online editing commands and interpreted the nature of those editing sessions that failed to produce a corrected or edited record. The results of the second phase of the study can be divided into two parts—one pertaining to generic findings and the other to findings that were unique to the system under study. In the first category, the researchers found that only a few sessions failed to produce one or more updates, that some commands are underutilized and may not need to be supported by the system, and that some library departments are much more likely than others to encounter editing problems. In the second category are minor "side-effects" of the system that can be repaired or, if not serious, ignored. However, even this second category has a generic lesson: only by undertaking a detailed examination of the editing performance of a staff in relation to the system's behavior can one uncover important and potentially troublesome man-machine interactions. ■■

PC-Mainframe Interface: An Innovative Use of Online Catalog Records for Automated Control of State Documents Collections

Margaret T. Mooney

California state government publications require cumbersome and time-consuming technical processing tasks at most California depository libraries. This unwieldy system occurs despite the relatively small number of titles received. Since the California State Library (CSL), which serves as the central agency for cataloging of California state documents, receives state publications at the same time as other depository libraries, depository items are distributed before the classification numbers are assigned to them by CSL. Libraries that choose to use these classification numbers as call numbers for their documents must therefore spend considerable effort searching subsequent issues of the printed California State Publications (CSP) to find matching cataloging records. In addition, acquisitions operations for titles that are not automatically received through the California depository program entail both tedious searching of CSP records against library holdings and labor-intensive work to prepare claim letters.

The recent inclusion of CSL's cataloging records in the University of California online catalog, MELVYL, has provided the U.C. Riverside Library with an opportunity to ameliorate the problem. This article describes an innovative solution that effectively utilizes CSL's cataloging records on MELVYL in the microcomputer environment.

AUTOMATION PROJECT

In the effort to streamline and enhance bibliographic control of its California documents collections, the Government Publications Department of the U.C. Riverside Library has created a microcomputer-based automated catalog for California documents from the University of California's MELVYL records. The project involves downloading the California State Library's MARC cataloging records from the MELVYL database and transferring these records to a dBASE file.

The resulting automated catalog, created without a single manual input, currently contains all California monographic documents cataloged by CSL since January 1988. The catalog has eliminated all tedious and time-consuming manual tasks previously associated with their processing. It serves both as an automated shelf list for U.C. Riverside's holdings of California documents and as a public-access catalog with a user-friendly, menu-driven interface that allows access to the database by various bibliographic elements. In addition, the catalog operates as an automated acquisitions tool and generates printed claim letters for titles that are not automatically received through the depository system.

PROJECT DETAILS

Downloading

The first step in the creation of this multipurpose catalog is the identification of all new California documents cataloged by CSL. Immediately after the monthly loading of the CSL tapes into the MELVYL catalog, staff members search the ADDED index with the command "F ADDED SINCE <DATE>AT CSL" to retrieve all new CSL cataloging records that were added from the last tape load. This strategy eliminates the need to search for individual titles. Moreover, the listing includes all of CSL's cataloging records for California monographic documents that would appear in the next issue of the printed CSP, not just the ones received by the U.C. Riverside Library. Thus, U.C. Riverside can use the database for acquisitions purposes.

The next step involves scanning the records by using the command "D TAG 086 CONT" to identify California documents. (The 086 field contains document numbers assigned by federal, state, and local agencies.) Approximately 100–200 titles appear each month, generally in a cluster. Staff then download these documents in the MARC display for-

Margaret T. Mooney is the Head of the Goverment Publications Department of the Rivera Library, University of California, Riverside.

MARC Tags	dBASE Fields
---	LOCATION
086	DOCNO
110/710	AGENCY
245	TITLE
260	IMPRINT
265	ADDRESS
300	PAGING
440/830	SERIES
500s	NOTES
---	HOLDINGS
500	STOCKNO
020	PRICE
915	AVCODE
910	CSP
901	CCSDNO
ID	MELVYLID
---	AGENCYNO

Figure 1. MELVYL MARC Tags and Corresponding dBASE Fields.

mat. A Zenith microcomputer with Hayes modem and Dialoglink communications software are used for downloading of records.

Transferring MELVYL Records to dBASE

A special program written by U.C. Riverside in the dBASE IV language transfers the downloaded records to a dBASE database. Figure 1 shows the structure of the dBASE file and the corresponding MELVYL record fields from which data are transferred.

Figure 2 presents a sample California state document in the MELVYL MARC display format and in the converted dBASE format.

BENEFITS OF THE AUTOMATED CATALOG

Shelf List

Once records are converted to the dBASE format, staff members compare each record in the database with depository documents received by U.C. Riverside. They then tag titles owned by the library in the "location" field of the records, making it an instant shelf list of the Library's California monographic documents collections. The database is indexed with dBASE indexing commands, and its records can be viewed in various orders (by classification number, title, agency, etc.) and accessed instantaneously by any bibliographic element.

MELVYL

Record

```
2.      ID 4797914 BASE   BL     STS n  REC am  ENC      DCF a     ENT 880414
           INT    REP    GOV s  CNF 0  FSC 0   INX 0    CTY cau   ILS
           MEI 1  FIC 0  BIO    MOD    CSC d   CON      LAN eng   PD 1988
020     $c $5.00  <BL,CSL>
040     C $c C  <BL,CSL>
043     n-us-ca  <BL,CSL>
086     L500.A52 1987 no.2 $2 cadocs  <BL,CSL>
110 10  California. $b Legislature. $b Senate. $b Select Committee on AIDS
        <BL,CSL>
245 10  Interim hearing on treatment of people with AIDS / $c California
        Legislature, Senate Select Committee on AIDS ; Senator Gary K. Hart,
        chair.  <BL,CSL>
260 0   Sacramento, CA : $b Joint Publications, $c [1988?]  <BL,CSL>
300     ii leaves, 63 p. ; $c 28 cm.  <BL,CSL>
500     Cover title: Treatment of people with AIDS.  <BL,CSL>
500     "State Building, 350 McAllister, room 1155, San Francisco, California,
        November 17, 1987, 1:00 p.m. - 5:00 p.m."  <BL,CSL>
500     Joint Publications Office stock no.: 324-S.  <BL,CSL>
650  0  AIDS (Disease) $z California  <BL,CSL>
740 01  Treatment of people with AIDS.  <BL,CSL>

CSL:    901 $aY$b88363$cCCSD     902 $a19880405151831.0     903 $aCSL
        904 $a19880405$b19880405     908 $aDH /BP     910 $acsp-88-04
        915 $aa
   Hol: 920 $aC        922 $aZCSL     924 $aC     926 $aMain
        930 $aL500.A52$b1987 no.2     935  1$a#33
```

dBASE

Record

```
LOCATION    D
DOCNO       L500.A52 1987 no.2
AGENCY      California. Legislature. Senate. Select Committee on AIDS.

TITLE       Interim hearing on treatment of people with AIDS / California Legisl
            ature, Senate Select Committee on AIDS ; Senator Gary K. Hart, chair

IMPRINT     Sacramento, CA : Joint Publications, [1988?]

ADDRESS     N
PAGING      ii leaves, 63 p.
SERIES

NOTES       Cover title: Treatment of people with AIDS.//"State Building, 350 Mc
            Allister, room 1155, San Francisco, California, November 17, 1987, 1
            :00 p.m. - 5:00 p.m.:
HOLDINGS
STOCKNO     324-S
PRICE       5.00
AVCODE      a
CSP         88-04
CCSDNO      88363
MELVYLID    4797914
AGENCYNO    L500
```

Figure 2. From MELVYL to dBASE—A Sample Record.

Public-Access Catalog

The automated catalog offers a menu-driven, public-access mode that allows the users to find documents by a given bibliographic element. Keyword access to the database, although slower in comparison to exact word searches, is also provided. The public-access mode is designed to prevent patrons from editing records, thus eliminating the danger of accidental changes or erasure of data.

Automated Acquisitions System

The automated catalog also generates claim

and order letters for titles not automatically received through depository distribution. In addition to printing bibliographic information pertinent to a title being claimed, the acquisitions program activates an agency address file that is linked to the automated catalog records by classification number. This linkage allows the printing of the appropriate address to where these claim letters may be sent. All tedious and time-consuming tasks involved in the manual typing of claim letters have been eliminated by this program.

FUTURE APPLICATIONS

In addition to the benefits mentioned above, the U.C. Riverside Library staff has foreseen potential applications in such areas as managerial analysis, development of collection profiles, and other projects related to the library's California documents collections. One of the most direct benefits of this project is the conversion of the library's shelf list for current California documents to a machine-readable form without mounting a special project. These machine-readable records will be matched against CSL's tape file as the University of California at Riverside Library prepares to provide complete cataloging information for its government publications collections. ■■

Sprint: A Microcomputer Program for Producing Call Number Labels

John K. Duke

Sprint is a locally designed program that prepares call number labels based on bibliographic records from the OCLC Micro Enhancer. The program prints preformated data above or below the call number. Sprint supports such features as multiple copies of labels, volume number increments, and options to print the call numbers or save them into a file. The program allows for significant savings in staff time.

Several years ago Virginia Commonwealth University (VCU) sought ways to improve the efficiency of its in-house marking and bindery support unit. Library staff was particularly interested in reducing the time needed to prepare call number labels. They hoped to eliminate the potential for errors inherent in a system dependent upon clerical staff retyping data transcribed by catalogers from the bibliographic record.

Staff members were spending twenty to thirty hours each week typing call numbers. When VCU began to streamline its cataloging operations by upgrading to OCLC M300 terminals and the Cataloging Micro Enhancer software,[1] the library investigated the production of spine labels. It looked into purchasing a system to automate call number printing, and also considered the call number formatting available through OCLC (the SL4 and SL6 commands). The available systems were rejected because they lacked flexibility and required too much editing to work the call numbers into VCU's preferred format. A large proportion of the processed materials for subunits of the library requires unique data above the call number.

The library decided to develop its own computer program. The goals of the program were:

1. To build in enough flexibility so that the library could reconfigure the system as needs changed without the aid of a programmer;
2. To derive all information directly from the bibliographic record, thus minimizing the need for operator intervention and increasing the accuracy of the product;
3. To permit manual input of individual call number labels in cases when the bibliographic record was not conveniently available or the label stock was nonstandard.

TECHNICAL DESCRIPTION

The library developed a program called "Sprint" to meet these objectives. Originally written in IBM's compiled BASIC, the program was later converted to QuickBASIC.[2] The program also incorporates routines from Key-1 Software[3] and Komputerwerk.[4] The

John K. Duke is the Interim Associated Director for Library Collections and Access Services, Virginia Commonwealth University, Richmond, Virginia.

program runs on a 256K IBM PC under DOS 2.x, using an eighty column monochrome monitor and one or two printers. It supports one or two disk drives and can also run from a hard disk. The program occupies about 131K of disk space.

The print format is vertical (i.e., call numbers print one on top of the other rather than stretching horizontally across the platen). The vertical format is designed to work with Gaylord's Sel-In machine. The library is on its second model of the machine. The first Gaylord machine used pressure on the platen and beneath the printer to stabilize the machine. The latest Gaylord model (290D) is a much more satisfactory device, but Gaylord must modify the platen by adding a gear as in the older style manual machines.

Although Sprint operates from files created with the Micro Enhancer, the automated printing requires conversion of the Micro Enhancer data files to ASCII format using a program known as BIBTODOS, which OCLC distributes as a separate utility.[5] Converting a disk containing 100 records takes about three minutes on an OCLC 80826 M310 computer with a hard disk.

OVERVIEW OF SYSTEM

Sprint has five modules, all accessible from a main menu (see figure 1). Although the first two modules are the workhorses of the system, describing the last three "definition modules" helps to clarify the structure of Sprint.

Define System

The system definitions configure the print settings and the disk drives. It is possible to attach two printers to the system. At VCU, one printer is permanently fitted with the Sel-In machine. A second printer generates sprocket-fed labels, reports, and individually fed, acid-free labels for special collection materials. Options for each printer include margin settings, printer control codes (such as for large or compressed print, letter-quality output, variable fonts, etc.), and spacing between labels. One may also define a printer for "variable"-length labels (such as the continuous Sel-In tape) or for "fixed"-length labels (where each label is die-cut to a standard size). File control tailors drives and paths to the system, per-

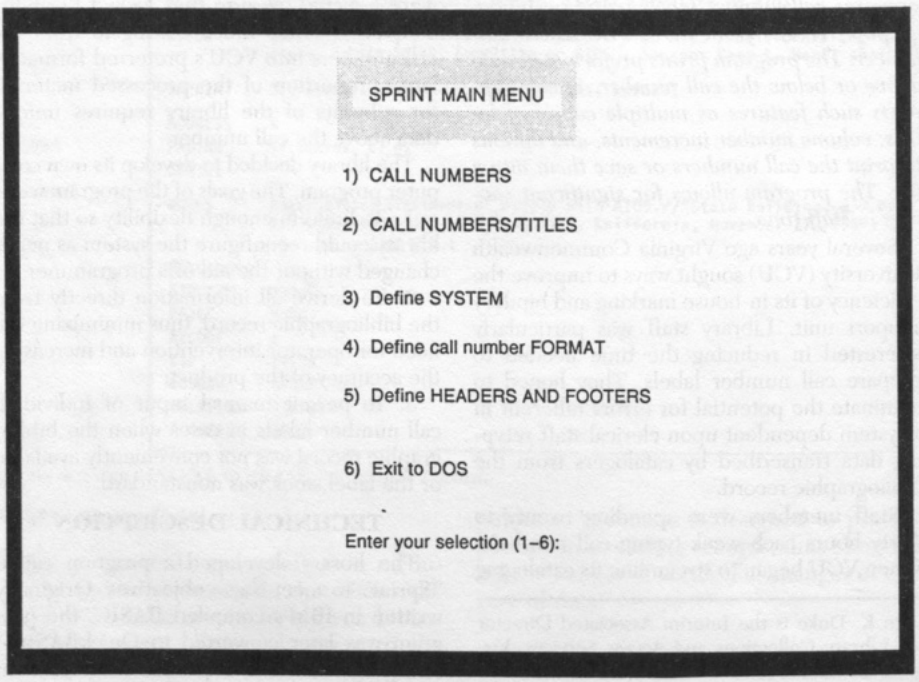


Figure 1. Sprint Main Menu.

mitting separate drives for the various files accessed or created by Sprint.

Define Call Number Format

The operator may control the classification structure portion of the LC number. The three styles include continuous alpha and numeric displays (AAnnn), alpha and numeric displays separated by a space (AA nnn), and alpha and numerics on separate lines.

Define Headers and Footers

The information that appears above a call number is the header, and information beneath the number is the footer. A header or footer may contain up to eight lines, although most are one or two lines. Characteristics of the bibliographic record trigger headers and footers. For example, the contents of one subfield "a" in an 049 field might print "REF" at the head of a call number, while another subfield "a" code prints "SPEC COLL" on separate lines.

Tags, indicators, subfields, or any other pieces of data within the variable fields can trigger a header or footer. Information can also combine from different parts of the record to signal Sprint to print a header or footer. For example, a code in a 049 field may only print the header if a particular subfield in a locally defined 949 field has a certain value. The operator saves defined headers and footers in a configuration file that loads when the program starts. Creating the header and footer records can be complex. At VCU, about eighty different headers are defined.

DATA PROCESSING

The operator initiates all data processing from one of two screens (see figure 2). The screens are identical in appearance and function except when processing multiple records. The data processing screens are rudimentary word processors that permit up to sixteen lines of text. This is useful for preparing labels for the occasional book, the nonstandard label (e.g., horizontal labels for sound cassettes and compact discs), books with multiple copies or volumes, or any other occasion when bibliographic records are not available in machine-readable form. Since both screens have independent configurations, VCU defines one with narrow margins for printing via the Sel-In machine running on an IBM Quietwriter

```
                              CALL NUMBERS

Start:          Stop:          Repeat:
    ....5....0....5....0....5....0....5....0....5....0....5....0....5....0....5....0....5
 1..........
 2..........
 3..........
 4..........
 5..........
 6..........
 7..........
 8..........
 9..........
10..........
11..........
12..........
13..........
14..........
15..........
16..........

    <F2>  Auto Print    <F3>  Header        <F4>  Footer
    <F5>  Copy          <F7>  Add           <F9>  Save        <F10>  Print
```

Figure 2. Data Processing Screen.

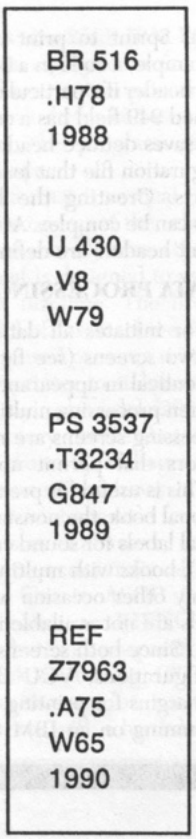


Figure 3. Call Number Processing Screen.

III; the other is designed for a second printer with wider defined margins for titles and call numbers in horizontal format. Sprint supports standard data entry features such as insert, delete, and typeover.

Function keys enable the following special features from the data processing screens:

Automatic Processing

The most powerful feature of the program processes OCLC Micro Enhancer records converted by the BIBTODOS utility. The operator may print the records or save them to a file.

The program recognizes call numbers in fields 099, 090, and 050. Sprint formats them according to the OCLC print program described in its format manuals. Logic imbedded within the program properly divides long numbers so that they print on separate lines. Sprint prints legitimate call numbers that otherwise generate OCLC error cards. If more than one call number field is on a record, the program selects the highest-numbered field. Defined headers and footers are added to the call number when appropriate.

The program handles automatic processing differently from each of the two screens. On one screen, it processes only the call numbers. This is the screen that is used for material prepared with Sel-In labels (see figure 3). A file of 100 records can be printed in about a minute and a half.

The second screen processes call numbers and titles. If the record has a personal author as a main entry (MARC field 100), the program adds the author's last name after the

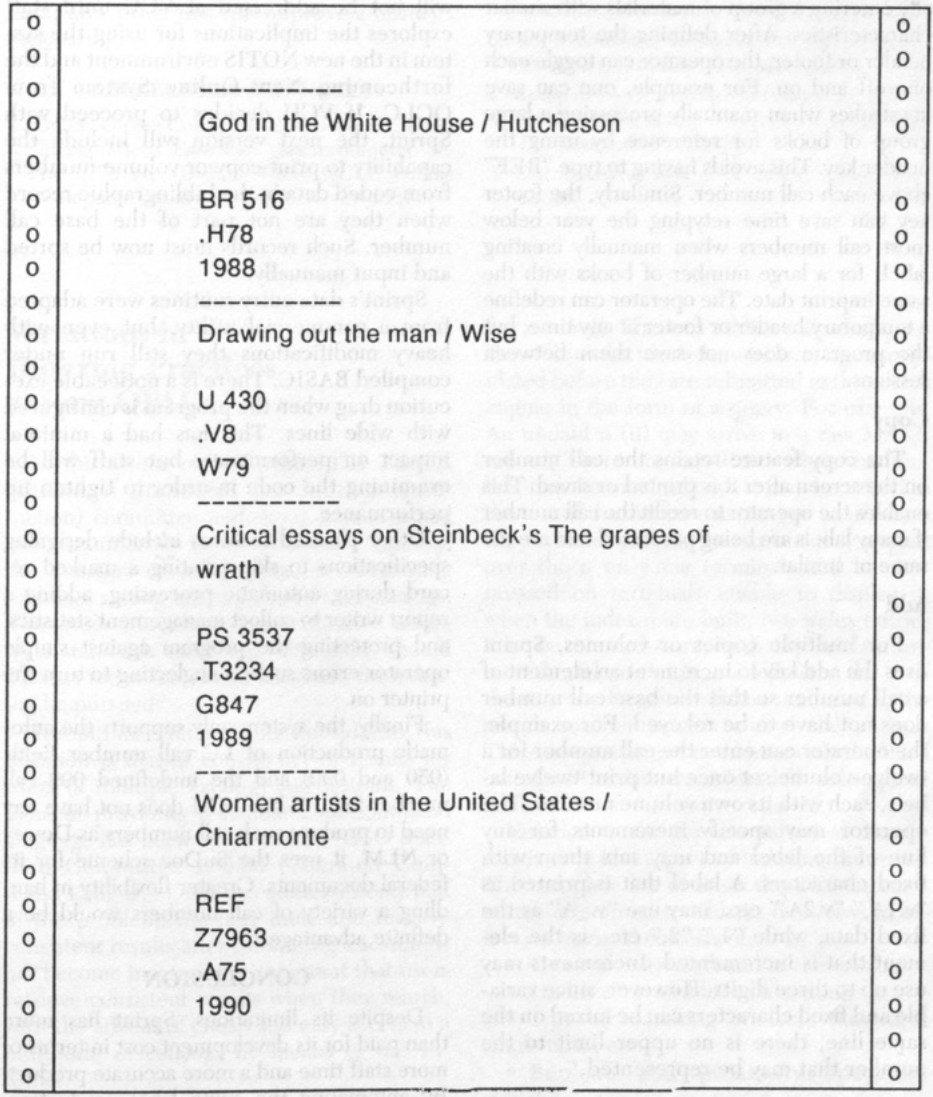


Figure 4. Call Number and Title Processing Screen.

title (see figure 4). The bindery preparation staff prints a hard copy of the file to create a list of volumes that were sent for binding. It is also used to verify the imprinted data when the volumes are returned. A sample of this file has been sent to a commercial binder to determine the feasibility of using it to imprint the title and call number directly onto the spine of the book.

Headers and Footers

In addition to defining and saving headers and footers through the configuration program, the operator can establish headers and

footers on the fly. This is useful when manually entering a group of materials with similar characteristics. After defining the temporary header or footer, the operator can toggle each one off and on. For example, one can save keystrokes when manually processing a large group of books for reference by using the header key. This avoids having to type "REF" above each call number. Similarly, the footer key can save time retyping the year below most call numbers when manually creating labels for a large number of books with the same imprint date. The operator can redefine a temporary header or footer at any time, but the program does not save them between sessions.

Copy

The copy feature retains the call number on the screen after it is printed or saved. This enables the operator to reedit the call number if many labels are being processed that are the same or similar.

Add

For multiple copies or volumes, Sprint uses the add key to increment an element of a call number so that the base call number does not have to be rekeyed. For example, the operator can enter the call number for a twelve-volume set once but print twelve labels, each with its own volume number. The operator may specify increments for any line of the label and may mix them with fixed characters. A label that is printed as "v.1A," "v.2A," etc., may use "v._A" as the fixed data, while "1," "2," etc., is the element that is incremented. Increments may use up to three digits. However, since variable and fixed characters can be mixed on the same line, there is no upper limit to the number that may be represented.

Save and Print

Call numbers (and/or titles), whether manually entered or automatically processed from an OCLC disk, may be routed either to a printer or to a file on a disk. Records are saved to disk in standard ASCII text format, although certain control characters separate data.

LIMITATIONS

Although Sprint has proven to be an enormously useful and powerful system for VCU, it has its limitations. The limitations will not be addressed at VCU until staff explores the implications for using the system in the new NOTIS environment and the forthcoming New Online System from OCLC. If VCU decides to proceed with Sprint, the next version will include the capability to print copy or volume numbers from coded data in the bibliographic record when they are not part of the base call number. Such records must now be sorted and input manually.

Sprint's data entry routines were adapted from a commercial utility, but even with heavy modifications they still run under compiled BASIC. There is a noticeable execution drag when the program is configured with wide lines. This has had a minimal impact on performance, but staff will be examining the code in order to tighten up performance.

Other planned features include designing specifications to skip printing a marked record during automatic processing, adding a report writer to collect management statistics, and protecting the program against simple operator errors such as neglecting to turn the printer on.

Finally, the system only supports the automatic production of LC call number fields (050 and 090) and the undefined 099 call number field. While VCU does not have any need to produce such call numbers as Dewey or NLM, it uses the SuDoc scheme for its federal documents. Greater flexibility in handling a variety of call numbers would be a definite advantage.

CONCLUSION

Despite its limitations, Sprint has more than paid for its development cost in terms of more staff time and a more accurate product. By automating the spine label production, VCU has been able to use clerks for other pressing projects.

As we evaluate the new NOTIS and OCLC environments, VCU may move away from the microcomputer as a tool for producing catalog records. Unless VCU staff develops microcomputer links with the mainframe to download records, it may be that Sprint will outlive its usefulness. In the meantime, however, the program continues to serve VCU well and it is an excellent

example of how locally developed microcomputer programs can improve a library's processing.

REFERENCES AND NOTES

1. OCLC Online Computer Library Center, *Cataloging Micro Enhancer* (Microcomputer Series), 2d ed. (Dublin, Ohio: OCLC Online Computer Library Center, 1985, rev. 1987).
2. Microsoft Corporation, *Microsoft QuickBASIC*, version 4.5 (Redmond, Wash.: Microsoft, 1988).
3. Key-1 Software, *Keytools2 Programming Aids* (Newport, R.I.: Key-1 Software, 1984).
4. Komputerwerk, *Finally!: A Subroutine Library for Compiled BASIC*, version 1.2 (Pittsburgh, Pa.: Komputerwerk, 1986).
5. Distributed as disk 3 of *OCLC Micro* (June 1986). ■■

Variations in Indexing Practices among OPACs

In the winter of 1989, the LITA TESLA (Technical Standards for Library Automation) committee addressed a topic it had been monitoring for several years; namely, the question of standardization of indexing practices among online public access catalogs. The issues were: how much variation exists, what standards exist, and, ultimately, should TESLA recommend that standards for indexing be pursued?

The ultimate rationale for considering this question is that of consistency of results for users of multiple catalogs and systems. The issue of providing a standardized command language has been addressed by NISO Standard Committee Z39.58 (Common Command Language), but this addresses only half of the problem: the user must also retrieve consistent results after submitting a search. It has become increasingly important that users retrieve consistent results when they search, with the advent of cheap and easy access to a multitude of catalogs by means of the Internet.

Variation in the representation of data in an online catalog (or in any text database) may occur in any of three places: (1) the storage of the raw text, (2) the display of the text, and (3) the indexing of those fields that are to be used for retrieval. There is additionally the question of how the user's search terms are manipulated before they are submitted to the search engine in the form of a query. For example: An umlaut *u* (ü) may arrive in a raw MARC record. There are four times when this umlaut may be manipulated: it may be stored in the text of the database in this raw form; when the text is displayed, the umlaut may appear over the *u* on some terminals and be suppressed on terminals unable to display it; when the indexes are built, two index entries may be created—one with the umlaut removed, and one with an *e* following the *u*. If the user tried to search for an umlauted *u*, the search scanner would remove the umlaut. This survey attempts to address only the third case: how the data are indexed for subsequent retrieval.

Variation in the representation of characters is only one aspect of indexing, however. The complete list of areas the TESLA committee attempted to address that have an effect on retrieval consists of:

- Types of material found in the catalog
- "Indexes" available for searching
- Type of searching available (keyword, whole term, etc.)
- Boolean and proximity operators available
- Presence of stopwords
- Normalization of nonalphanumeric characters (diacritics, special characters, and punctuation)
- Normalization of text (e.g., changing to all upper case)
- Special treatment in the indexing of initialisms, acronyms, and abbreviations
- Authority control
- MARC fields and subfields included in each index

This survey was carried out by a subcommittee of the LITA TESLA committee. Its members were Katharina Klemperer, Jim Michael, and Sara Randall. This report was prepared by Katharina Klemperer, who is director of library automation at Dartmouth College.

Clearly, variations among systems in any of these areas will cause variations in retrieval, no matter how standardized their user interfaces or command languages.

PROCEDURE AND SURVEY INSTRUMENT

Over a period of months, a TESLA subcommittee developed a survey instrument that would cover the most common areas of variation in indexing. Naturally the first version was far too long and complex. After subjecting several colleagues to a pretest and incorporating their suggestions, the committee distributed a final version to approximately sixty vendors and libraries, including suppliers of all major automated OPACs.

Microcomputer- and CD-ROM-based systems were not included in order to keep the project manageable. The conclusions may neverthless be applied to these kinds of systems. The mix included both commercial products and locally developed systems.

Respondents were asked to adhere to the following guidelines:

- To answer only with regard to those features that are actually implemented and available, and
- To answer only with regard to the public access catalog part of the system, not circulation, technical services, or other staff-only functions.

The first round of questionnaires was distributed in the fall of 1989, with follow-up messages during 1990. By August 1990, thirty-two responses had been received. The results were tabulated and summarized but were not subjected to any rigorous statistical analysis.

Twenty of the respondents represented systems that are available commercially; the remainder were locally developed systems that are not available for purchase. Most of the systems represented in this analysis are geared toward large academic research libraries, although public library systems are included. The survey instrument is reproduced in figure 1.

RESULTS AND CONCLUSIONS

The summarized results have been entered manually into the survey instrument in figure 1. Totals will not always add up to thirty-two because some respondents did not answer all questions.

A number of conclusions may be drawn from the results. Not surprisingly, there is indeed a great deal of variation in the practices among the major OPACs in use today. Section 1 attempted to cover general indexing practices. A random list of items of interest follows:

- A surprising number of OPACs do not include all material types (question 1).
- Very few systems allow free-text searching of the entire bibliographic record. This is recorded in question 2 as a "general" search.
- Nearly all systems provide keyword and Boolean searching, or claim that they do (questions 4–5).
- Very few provide adjacency or phrase searching (question 5a).
- There is an enormous variation in the treatment of punctuation, diacritics, and special characters. For many characters, there is not even a predominant translation. This will cause users to retrieve widely different results on different systems (question 7).
- The use of stopwords varies widely. For commercial systems, these are usually configurable by the customer (question 6).

Section 2 of the instrument was intended to compare the actual MARC data elements that are indexed in various systems. In response to this section, the investigators received a substantial number of detailed responses. It would be possible to tabulate these responses, tag by tag and subfield by subfield. The resulting tabulation would indicate that there is indeed significant variation in the MARC indexing for author, title, and subject searching among systems.

A not insignificant outcome of this survey was a wealth of data that could be used in evaluating systems for purchase. The committee feels that the survey instrument itself might be helpful as a guideline for reviewing systems when making a selection. Because this was not the original intent of the survey, however, the individual responses will not be made public.

Other conclusions can be drawn about the nature of the documentation that different systems provide. Some respondents sent tables or sections of programs (which they were encouraged to do if it made it easier to supply a response), others submit-

PLEASE TELL US HOW MUCH TIME YOU SPENT FILLING OUT THIS SURVEY:_____

What is the name of the system being described?_____

If this system is based on a commercial system with a different name, what is the name of that system?_____

What is the name of the responding institution?_____

Name of contact person for this questionnaire:_____

Mailing Address:_____

Phone:_____ e-mail address:_____ Fax:_____

Section 1 - General Database Information

1. What types of material does this database include? (circle all that apply)

books	32	manuscripts	25	maps	28	music scores	28
serials	30	datafiles	26	AV	30	sound recordings	30

Other_____

2. List the indexes that can be used to initiate a search:

GENERAL - 8 SERIES - 17 NOTES - 6
AUTHOR - 30 ISBN/ISSN - 22 OTHER TEXT - 17
TITLE - 30 LCCN - 17
SUBJECT - 29 CALL NUM - 18 OTHER NUMBER - 18

3. List the "indexes" or data elements (for example, date, or language) that can be used to limit or modify a completed search, but not to initiate one:

DATE - 18 MATERIAL TYPE - 15
LOCATION - 11
LANGUAGE - 16 OTHER - 6

4. Check what kind of search can be performed in this system(more than one may apply):

29 keyword (each word can be searched individually or in combination)
29 whole or exact term search (the whole field is entered by the user)
32 truncated whole or exact term search (the system retrieves whole fields that begin with the string entered by the user)
12 other (describe)_____

5a. If your indexing supports keyword searching, what Boolean or proximity operators can be entered explicitly by the user? (circle all that apply)

25	AND	21	NOT
21	INCLUSIVE OR	3	ADJACENT IN EITHER ORDER
5	EXCLUSIVE OR	8	ADJACENT IN SAME ORDER
4	none	5	OTHER _____

Figure 1. Survey of Online Access Catalog Indexing—LITA TESLA Committee, October 1989.

ted pages from their manuals, and still others indicated that the information was too difficult to ferret out of their programs. Again, this sort of conclusion was not in the original scope of the project, but an awareness of it could aid customers in selecting a system.

RECOMMENDATION

The TESLA committee feels that a great deal of valuable information has been gathered from this survey and would like to suggest the development of indexing guidelines in certain areas, particularly the treat-

5b. If keyword searching is available, what Boolean or proximity operator is assumed between two adjacent keywords (AND, OR, ADJACENCY, etc.)? AND-18, ADJ-4, OR-1
2 Configurable by customer

5c. Can keywords be truncated? Yes 24 No 4

6a. Does your catalog use a stopword list? Yes 29 No 2

6b. If Yes, list the stopwords and the indexes to which they apply:

There was enormous variety here.
Many are configurable by customer.

6c. If details of stopword are not known, approximately how many words appear on it?_____

6d. If this system can be exported, can the stopword list be customized for each installation? Yes ____ No ____

7. Indicate how the general indexing routine treats each of the following punctuation, diacritics and special characters, by circling the treatment you use. ("Lv" means leave as it appears, "Rem" means "remove", "Sp" means "translate to space", "Trans to" means "translate to") Note that this question refers to *indexing*, not display.

period (.)	Lv 4	Rem 9	Sp 11	Trans to.	Other 2
hyphen (-)	Lv 7	Rem 3	Sp 13	Trans to.	Other 4
dash (--)	Lv 1	Rem 10	Sp 11	Trans to.	Other 2
slash (/)	Lv 3	Rem 8	Sp 14	Trans to.	Other 1
comma (,)	Lv 3	Rem 9	Sp 11	Trans to.	Other 2
underscore (_)	Lv 0	Rem 15	Sp 9	Trans to.	Other
plus sign (+)	Lv 1	Rem 14	Sp 10	Trans to.	Other
exclamation (!)	Lv 2	Rem 11	Sp 12	Trans to.	Other
at sign (@)	Lv 1	Rem 13	Sp 11	Trans to.	Other
question (?)	Lv 2	Rem 12	Sp 12	Trans to.	Other
sharp (#)	Lv 2	Rem 13	Sp 8	Trans to.	Other
ampersand (&)	Lv 5	Rem 9	Sp 7	Trans to. AND-5	Other 1
dollar sign ($)	Lv 1	Rem 13	Sp 11	Trans to.	Other
parentheses	Lv 3	Rem 10	Sp 11	Trans to.	Other
apostrophe (')	Lv 5	Rem 13	Sp 3	Trans to.	Other 3
double quote (")	Lv 1	Rem 15	Sp 10	Trans to.	Other
asterisk (*)	Lv 1	Rem 14	Sp 11	Trans to.	Other
less than (<)	Lv 1	Rem 14	Sp 10	Trans to.	Other
greater than (>)	Lv 1	Rem 14	Sp 10	Trans to.	Other
equal sign (=)	Lv 1	Rem 14	Sp 12	Trans to.	Other
brackets ([ ])	Lv 2	Rem 13	Sp 9	Trans to.	Other
braces ({ })	Lv 1	Rem 15	Sp 9	Trans to.	Other
backslash (\)	Lv 2	Rem 13	Sp 10	Trans to.	Other
vertical bar (|)	Lv 1	Rem 15	Sp 6	Trans to.	Other
alpha (a)	Lv 1	Rem 8	Sp 0	Trans to. A-11, ALPHA-6	Other
beta (ß)	Lv 1	Rem 8	Sp 0	Trans to. B-11, BETA-6	Other
subscripts	Lv 1	Rem 8	Sp 1	Trans to. numeral-14	Other
superscripts	Lv 1	Rem 8	Sp 1	Trans to. numeral-14	Other
degree (°)	Lv 1	Rem 15	Sp 4	Trans to.	Other 1
tilde (~)	Lv 0	Rem 22	Sp 3	Trans to.	Other
script el	Lv 1	Rem 6	Sp 0	Trans to. L-14, EL-1	Other
musical flat	Lv 1	Rem 13	Sp 1	Trans to. B-4, FLAT-4	Other
umlaut (¨)	Lv 1	Rem 22	Sp 0	Trans to. E-1	Other 1

Figure 1. Survey of Online Access Catalog Indexing—LITA TESLA Committee, October 1989 (cont.).

c cedilla (ç)	Lv 1	Rem 10	Sp 0	Trans to C-12	Other
angstrom a (å)	Lv 1	Rem 11	Sp 0	Trans to A-11, AA-1	Other 1
æ digraph	Lv 1	Rem 4	Sp 0	Trans to AE-21, A-1	Other
œ digraph	Lv 1	Rem 4	Sp 0	Trans to OE-21, O-1	Other
cross d	Lv 1	Rem 4	Sp 0	Trans to D-22	Other
eth	Lv 1	Rem 4	Sp 1	Trans to D-12, TH-6, :-1	Other
Turkish dotless i	Lv 1	Rem 4	Sp 1	Trans to I-21	Other
polish el	Lv 1	Rem 5	Sp 0	Trans to L-21	Other
hook o	Lv 1	Rem 4	Sp 0	Trans to O-21, U-1	Other
hook u	Lv 1	Rem 4	Sp 0	Trans to U-21, Y-1	Other
slash o (ø)	Lv 1	Rem 4	Sp 0	Trans to O-21, OE-1	Other
thorn	Lv 1	Rem 4	Sp 0	Trans to TH-18, 8-1, :-1	Other

Other:

10 made other translations	Lv	Rem	Sp	Trans to	Other
	Lv	Rem	Sp	Trans to	Other
	Lv	Rem	Sp	Trans to	Other
	Lv	Rem	Sp	Trans to	Other
	Lv	Rem	Sp	Trans to	Other

8. List any abbreviations which are automatically expanded for indexing:

Abbreviation	Expansion(s) that are indexed
	7 systems did some automatic abbreviation expansion. 2 had configurable lists.

9. Describe general normalization applied to all basic indexes. (For example, all alphabetics are changed to upper case)

All systems changed alphabetics to upper case
Many removed leading, trailing and multiple blanks
2 did extensive normalization

10. Describe any special treatment given to initialisms and acronyms.

6 removed periods separating single letters
1 removed spaces separating single letters
2 were customer-configurable

11. Describe any special indexing treatment given to numbers.

Some normalized and zero-padded numbers.
Some modified Call numbers.
1 changed Roman to Arabic numerals.

12. Authority control normalization:

a. Does the system allow direct retrieval of bibliographi records if the user searches for a non-standard form (i.e., on a cross-reference?) Yes 18 No 14

b. If Yes, to which indexes does this apply?

Usually author, uniform title, series and subject.

e. If there is an authority file, can it be searched independently? Yes 21 No 8

f. If Yes, what authority indexes can be searched? (circle all that apply)

Personal name	20	Conference name	19	Series	16
Corporate name	19	Uniform title	17	Subject	19
Other MeSH					

Figure 1. Survey of Online Access Catalog Indexing—LITA TESLA Committee, October 1989 (cont.).

Section 2 - MARC fields indexed

2. Please fill in the MARC *bibliographic* fields (tags) and subfields that are indexed for each of the following indexes:

AUTHOR

tag	subfields	tag	subfields

TITLE

tag	subfields	tag	subfields

SUBJECT

tag	subfields	tag	subfields

Figure 1. Survey of Online Access Catalog Indexing—LITA TESLA Committee, October 1989 (cont.).

ment of special characters, diacritics, and punctuation; the handling of abbreviations; and the fields and subfields to be included in each index. These would aid system designers in developing their translation tables and would provide a service to users.

However, the committee feels that, despite the variation in indexing practices and the potential advantage of an actual standard for indexing of OPACs, it would not be wise to

develop an actual standard, for two largely practical reasons.

First, the index structure is such an integral feature of a system that it cannot be modified without redesigning the entire system. Unlike a standard command language, which can be superimposed upon an existing system and which can exist side-by-side with a "native" language, an indexing standard would require the rebuilding of entire databases. Second, there are expenses involved in providing excellent indexing and access. Keyword access with a full range of Boolean and proximity operators may be ideal, but it requires more disk space and a more complex search engine than a simple whole-term search. Similarly, the capability of retrieving headings by searching for cross-references is an admirable goal, but it adds tremendous overhead to the development costs of the system. It is cheaper and easier to provide simpler indexing, and customers should be able to choose a lower quality for a lower price. Imposing standards in such areas is counter to the principles of a market economy. In general terms, a standard that covered these areas would clearly never be approved or followed, and proposing one would be inappropriate.

The TESLA committee is submitting this recommendation to the LITA board of directors in the hopes that it will make an official recommendation to the national standards-developing bodies. ■■

Recent Publications

Book Reviews

101 Uses of dBASE in Libraries. Supplement to Computers in Libraries, no.12. Ed. by Lynne Hayman. Westport, Conn.: Meckler. 1990. 118p. $34.95 (ISBN 0-88736-427-6).

If there really were 101 uses here, this slim volume would be a bargain at about thirty-five cents per use. However, this is not the annotated inventory of uses for dBASE software that the title leads us to expect. Its papers describe only seven distinct applications: document delivery, records management (two articles), periodicals listing, circulation, newspaper indexing, serials management, and acquisitions. An appendix contains two non-dBASE applications: an archive inventory system using a dBASE-like program called VPInfo and an approach to direct mail fundraising using the program PowerBase.

Editor Lynne Hayman does a good job of defining relational databases and the general features of dBASE, but she does not introduce, situate, or even comment upon the articles that follow except to say that the volume's intention is "not to share programming code . . . but to illustrate the wide range of library functions which are well served by locally designed applications of database management systems." Hayman tries to save the integrity of her title, but she only manages to list twenty-eight "uses." The bibliography reveals a number of articles on dBASE II as well as four pages of general articles, many of which are industry reports from such publications as *InfoWorld, Byte,* and *PC Magazine.*

Anne F. Corbett and Valerie M. Bennett report in separate articles on a system at the Library of the College of Physicians of Philadelphia that tracks document delivery transactions. The way in which the system was developed illustrates how to grow a local system. Bennett's treatment is more detailed and might have sufficed by itself.

This is less true elsewhere. Russell Miller demonstrates the first steps in discovering the use and power of dBASE, but his periodical listing system can hardly be considered a model. Nick Carter's treatment of a circulation system for a technical library illustrates the troubles that come from "cloning" an outdated batch processing system. The serials management article by Betty Johnson and Linda Grooms is intriguing but rather elliptical. When Francisco Garcia-Ayvens discusses newspaper indexing, he shares an ingenious solution for using multiple subject headings for each bibliographic citation, offering to share his code with correspondents. His solution, however, requires a great deal of user-intervention that should be handled by the dBASE programs.

Most helpful of this uneven batch is Robert P. Rynkiewicz's article on Atlantic City Public Library's acquisitions system. Rynkiewicz begins with a skeletal system described in Karl Beiser's *Essential Guide to dBASE III in Libraries* and adapts it to the real world of a multibranch public library.

While there is a bit of help and wisdom here, the book lacks the substance one expects in a publication devoted to this widely used database management software. A more useful publication covering dBASE applications is Karl Beiser's *Guide,* which provides a more thorough basic overview.—*John Blegen, Glenview Public Library.* ■■

Crawford, Walt. *Desktop Publishing for Librarians.* Professional Librarian Series. Boston: Hall, 1990. 370p. $38.50 (ISBN 0-8161- 1929-5).

This source provides useful information on desktop publishing regardless of one's level of expertise. It is logically organized and allows for quick reference to information on any particular topic. Examples are well chosen and add clarity, especially for the beginner. While full-screen examples are sometimes difficult to read, the opportunity to view all

the information afforded by the full-screen display is helpful. Coverage is well balanced, offering a good introduction to graphic design as well as computer hardware and software packages. The author offers advice on when not to use desktop publishing and provides warnings of potential pitfalls.

The book is divided into three main sections: uses for desktop publishing, which serves as an overview; document design and production, which focuses on layout; and tools for desktop publishing, which discusses hardware, software, and graphics. A number of library-related projects are discussed, including signs, bibliographies, brochures, annual reports, and newsletters. Actual library documents are reproduced as examples. The extensive glossary is helpful to beginners. The chapter title "Bibliographic Essay" offers reviews of more than a dozen recent books and periodicals in this area. As a frequent reviewer of books in this field, the author demonstrates his special expertise in the survey of desktop publishing literature. Crawford also offers his personal opinion and gives mostly excellent advice. He shows a clear bias toward IBM and compatible products, perhaps because libraries seem inclined to do so. The book seems to advocate the centralization of desktop publishing operations within the library or within departments, a recommendation with which some may disagree.

Overall, the text is extremely readable. The author should be applauded for frequently achieving clarity with cleverness. In some instances, however, this clever, chatty style implies an arrogance that may offend desktop novices ("a good drawing program . . . might even allow a graphics illiterate to produce something worthwhile" (p.331)). An excellent feature is the "Tips and Reminders" section at the end of each chapter. These are written in a concise, direct style and are useful pointers that apply both common sense and experience.

The author is very knowledgeable about desktop publishing and has some helpful advice, especially for beginners. His book is one of the clearest that I have read on the subject.—*Connie V. Dowell, University of California, Santa Barbara.* ■■

Dewey, Patrick R. *FAX for Libraries.* Supplements to Computers in Libraries, no.13. Westport, Conn.: Meckler, 1990. 128p. $39.95 (ISBN 0-88736-480-2).

About 60 of the 128 pages in this book consist of a directory of telefacsimile (fax) machines available at the time the book was written. Each entry states the make, model number, and a few pertinent technical facts such as group compatibility, transmission time, and paper requirements. A price is quoted for some, but not all, models described. The directory contains no data regarding service record, reliability, ease of use, or other consumer information that might influence selection.

As an introduction to the directory, the author has provided a very brief overview of the history of fax technology and a few paragraphs about early use of fax in libraries (covering the period from 1983 to 1985). Two chapters address selection criteria, management issues, and alternatives to fax. The most important selection criteria are price and reliability; modems and fax boards installed in a PC may be an appropriate alternative for standalone fax in some environments.

The selected bibliography (seventeen articles published between 1984 and April 1989) is useful to illustrate the growing acceptance of fax in libraries but is of no help to the librarian who wants to buy a good, reliable fax machine today. For that kind of information, a few minutes with any of the online or CD-ROM periodical indexes will reveal a plethora of up-to-date consumer information articles. Libraries don't really need this book, particularly one priced at $40. The preliminary information about fax technology and library applications is too brief and superficial for a monograph, and the marketplace for electronic gadgetry is so volatile that its product information will probably be out of date by the time you read this review.—*Harold D. Neikirk, Western Maryland College.* ■■

Johnson, Richard D., and Harriet H. Johnson. *The Macintosh Press: Desktop Publishing for Libraries.* Supplements to Small Computers in Libraries, no.9. Westport, Conn.: Meckler, 1989. 180p. paper, $39.50 (ISBN 0-88736-287-7).

Richard and Harriet Johnson have produced a book for those who have heard of desktop publishing but have never considered attempting it themselves. The Macintosh

Press seeks to persuade beginning and intermediate computer users that the user-friendly Macintosh PC is the perfect machine for launching their desktop publishing careers. Experience with a Macintosh is recommended but not necessary.

The Johnsons begin with a brief discussion of the history of graphic presentation in libraries, beginning with the hand-printed sign, to the typewritten flyer, and ultimately to the computer- produced newsletter. The remainder of the book covers options for planning, developing, and producing a publication on a Macintosh computer.

Coverage is broad but lacks depth, and the presentation is somewhat unsophisticated in comparison with Peter Stubley's *Desktop Publishing for Librarians on the Apple Macintosh*, which was reviewed in *ITAL* September 1990. The Johnsons present type styles, fonts, and use of graphics, and they discuss familiar software packages for word processing and graphics. A unique feature is the discussion of spreadsheets and database programs in library publications. In addition, they compare the advantages of three well-known page makeup software programs.

The illustrations and examples from publications the Johnsons have produced themselves are used effectively. The book concludes with a bibliography arranged by topic, a list of software producers, and an index.

Since more detailed "how to" manuals will be needed to become truly proficient, this book is only the beginning for a would-be desktop publisher. However, *The Macintosh Press* makes a welcome contribution by showing that those with modest word processing experience can be successful in desktop publishing.—*Carol Caro, Boston College.* ■■

Library and Information Technology Standards. Ed. by Michael Gorman. LITA Monographs, no.1. Chicago: Library and Information Technology Association, ALA, 1990. 90p. paper, $18.50; LITA members $16.65 (ISBN 0-8389-7431-7).

This book contains the papers presented at two technical sessions dealing specifically with library information system standards held during LITA's second national conference in Boston in October 1988. Like so many conference proceedings, the work is valuable because it expands the audience beyond the original attendees. It is marred, however, by the passage of two years without the inclusion of updated information.

Michael Gorman, the editor, follows the format of the original two sessions by presenting the eight papers under two main sections, "Developing Standards" and "Implementing Standards." There is also an introduction by Marilyn Nasatir, former chair of the Technical Standards for Library Automation committee (TESLA), which hosted the sessions at the conference. Her essay outlines the need for awareness of standards issues and for involvement by the user community.

Pat Harris, executive director of the National Information Standards Organization (NISO), begins the first section with a paper that provides a concise overview of the three stages of standards development. This includes a description of the difference between the American National Standards Institute (ANSI) and NISO organizations and some tidbits concerning the origins of the term Z39.

Katharina Klemperer follows with a description of the work of NISO Committee G, charged with developing the Common Command Language. The author provides valuable historical information about a standard that has been in development for a decade. Unfortunately, the reader is left to look elsewhere for information on the draft standard during the past two years.

Nolan Pope introduces ALA's Machine Readable Bibliographic Information (MARBI) Committee, guiding the reader through the evolution of the USMARC Format for Holdings and Locations. As a member of the original holdings committee, he is well equipped to cover this area.

Tony Carbo Bearman concludes the "Developing Standards" section with a description and detailed syllabus of a model standards curriculum for library and information sciences. The syllabus alone educates the reader in sorting out standards.

The second section of the book, "Implementing Standards," includes three papers that describe the author's experiences while conforming to and, in some cases, not conforming to national standards. Standardization can pose both advantages and disadvantages. The advantages may be economic. They may foster improved performance or ease of

use, or they may promote a higher level of quality. Disadvantages of standardization include application of inappropriate or unnecessary standards or simply bad timing. Gary McCone describes the National Agricultural Library's experience in following the USMARC family of formats. Naomi Ronen provides the perspective of the law librarian. Steve Salmon reports on aspects of standards issues that arise in circulation system development.

In the final essay, Walt Crawford talks about "Un-standards," technical standards that are abandoned, not approved, or approved but not widely adopted. While this is fun reading, it also makes a very important point: "We do not need standards for everything, and we do need to avoid bad standards."

Each paper essentially stands alone; together the papers provide a composite library and information standards picture that, although not global as the title would suggest, offers a useful introduction to current standards issues and terminology.—*Sylvia MacKinnon Carson, Pennsylvania State University.* ■■

Library Education and Employer Expectations. Ed. by E. Dale Cluff. Binghamton, N.Y.: Haworth, 1990. $32.95 (ISBN 0-86656-896-4).

One of the major issues facing the library profession is the identification of the kinds of librarians and support staff who will be needed in the twenty-first century. Discussion must focus on where this staff will come from and where and how they will be prepared for their work in libraries. Dialogue must engage both library educators and the profession as a whole. This collection is an interesting starting point for reviewing some employer perspectives on the people that libraries should employ and the knowledge and skills needed by professional staff that can be developed in library education programs.

Cluff organizes his collection of nineteen essays around five perspectives of library education and employer expectations, beginning with accreditation as presented by Herman Totten, immediate past-chair of ALA's Committee on Accreditation. The remaining four perspectives include those of two library school deans, employers in a variety of libraries, recent graduates of library schools, and a doctoral candidate, who shares the results of a study that attempts to identify the policies and attitudes of academic libraries concerning the value of an ALA-accredited M.L.S. degree when recruiting librarians.

There are no real surprises in the essays, some of which are based on research. Most are opinion pieces or reports on surveys of the literature. The majority of the contributors emphasize the need for librarians who can ask questions, think critically, solve problems, relate to technology and apply it toward library functions, and demonstrate managerial knowledge and abilities. Several of the writers value candidates who are risk takers. One must note, however, that a persistent complaint of graduate students is that they land in positions where questioning the status quo is not looked upon favorably.

Only Gary Strong, a California state librarian, emphasizes passion, compassion, and love of books as characteristics to be desired by employers. Both Rosemary Martin, who addresses public library employer expectations, and Strong make emphatic pleas for people-oriented staff. Human issues would seem to be a critical factor in planning library service in the next century when library work will be increasingly dominated by technology. Many of the essayists continue to stress the need for preparation programs that provide a prudent balance between theory and practice. Some recommend increased use of internships for both students and faculty.

The collection is weighted with discussions of the academic library, its needs and perspective. The only special libraries discussed are those that serve the medical, law, archival services, and science and engineering professions such as the National Library of Agriculture. Authors of all but the latter appear to be affiliated with academic institutions. Another interesting omission is any direct reference to one of the profession's three two-year library education programs. Many references are made in the essays to these programs, and a fine essay by Ronald Rodriguez, perhaps one of the best in the collection, discusses his experience in the UCLA program in "Library School Education in Retrospect: One Latino Graduate's View." Personnel issues facing 90,000 school libraries caught in a mesh of economic crisis, reduction of state certification requirements, technological change, and

a paucity of ALA-accredited library education programs across the nation are inadequately represented in Polk and Kahler's essay, which reports on the results of a survey of Texas school library media specialists and their employers. Rosemary Martin briefly addresses the deprofessionalization and the shortage of children's librarians in public libraries, the only essay on public libraries.

Even though this volume, like much of library literature, generally uses the academic library as its benchmark, it will prove useful to library school students and faculty interested in employment problems and employer perspectives. Librarians concerned with recruitment practices will also be interested. This book may motivate more discussion in the profession.—*Marilyn L. Miller, University of North Carolina, Greensboro.* ■■

Modern Copyright Fundamentals. Key Writings on Technological and Other Issues. Rev. ed. Ed. by Ben H. Weil and Barbara Friedman Plansky. Medford, N.J.: Learned Information for the American Society for Information Science, 1989. 460p. $39.50 (ISBN 0-938734-33- 4).

Modern Copyright Fundamentals first appeared in 1985 as a collection of seventy-five articles from 1974 to 1984 on copyright law and the complexities of applying it in an era of rapid technological change. The articles were photographically reproduced and organized into nine chapters prefaced by commentaries by the authors: Overview and Basics, Creators and Owners, Users, Intermediaries/Information Services, Basic Issues, Technological Issues, Legislative and Judicial Issues, International Aspects, and Future Developments. In this 1989 reprint for ASIS, a four-page update plus four pages of bibliographic citations (to 1988) have been added by the authors. Ben H. Weil continues to publish on copyright issues, though now retired from Exxon Research and the board of the Copyright Clearance Center, which he "designed and put into action." Barbara Friedman Polansky has been copyright administrator for the American Chemical Society's Books and Journals Division since 1979.

For those who did not purchase the 1985 edition, the reprinted collection brings together key articles on a historical period that saw the application of the 1976 copyright revision to photocopying, a definition of fair use, resolution of the Betamax case concerning video copying, and the formation of the Copyright Clearance Center. Some sections on withdrawn legislation have been superseded (Umbrella Statute Amendment, p.332–38; Manufacturing Clause, p.380–93) and International Aspects have been altered by the passage of the Berne Convention, ratified in 1989.

Many of the technological issues that are discussed in the articles are still of vital importance to libraries today—ownership of electronic media, downloading, electronic document delivery, copyright of computer programs and chips, international distribution of database information—but the revised edition does not address developments that change the dimension and nature of technological issues. CD-ROMs and large optical and magnetic disk storage capacity have made downloading large files commonplace; networks from LANs (local area networks) to the Internet to the proposed NREN (National Research and Education Network) facilitate data transfer and make copyright difficult to police. Additionally, a national deposit of machine-readable information is under consideration, fax has invaded all arenas of information transfer, and software copyright is a major issue. Those who make their way through the book's ten years of response of copyright law to technology will need to use the updated bibliography as well as their own library collections to bring the discussion into the nineties.

The book's final section, Future Developments, includes thoughtful consideration of the viability of a copyright law that originated for the technology of the printing press at a time when centralized control of the production process was possible. Articles debate the survival of such a law in an era of new technology that undermines the definition of intellectual property as well as the means of protecting such property in an environment of electronic use.—*Mary Kay Duggan, University of California, Berkeley.* ■■

Warwick, Robert T., and Patricia E. Jensen. *Using OCLC: A How-To-Do-It Manual for Libraries.* New York: Neal-Schuman, 1990. 145p. (ISBN 1-55570-037-3).

This manual is intended for use by library science students and by library practitioners. The growth and availability of regional and national bibliographic utilities make it necessary that those seeking employment or those employed in libraries have some experience with bibliographic searching and database development.

The authors have taken a simple, basic approach to learning and using OCLC. Information is presented in a logical sequence, allowing for self-directed study. Readers are taken through the fundamental steps of communication with the system in Part 1. Unfortunately, the approach is so basic that the only advice offered for dealing with terminals or PC workstations is that the user "familiarize himself with the hardware at hand before beginning."

Part 2 deals with searching the system, including the construction of search keys, displaying search results, and qualifying searches. Part 3 offers an extensive review of the cataloging subsystem. Chapters deal with the appearance of bibliographic records; record preparation, including tagging; the input process; editing records; and finally, some information about saving records, canceling a transaction, OCLC's profiling procedures, and completing a transaction.

All of this is presented logically and in a sequential manner. One drawback is that once OCLC's new online system is put into general use, this manual will have limited applicability. In the meantime, this manual leaves nothing unsaid, with examples taken directly from the online system. Additionally, this manual is also a workbook. Each segment is accompanied by exercises meant to be used with the online system for practice and reinforcement. This adds to its value as a training tool.

Another useful feature is the format of the pages in the manual. There is a generous amount of blank space on each page, where notes may be written by the user relative to the material being discussed. Writers of other manuals would do well to follow the authors' example.

There are useful appendixes. Summaries of search keys and qualifiers are given. Included are stop list words, sample displays, display and editing commands, work forms, and an abridged summary of the MARC format. References, a glossary, and answers to the exercises complete this seventh in a series of how-to-do-it manuals for libraries. It should be placed at the OCLC workstation for easy reference or made part of the library school's reference or reserve collection.—*John A. Richardson, New York Public Library.* ■■

Software Reviews

PC Tools Deluxe version 6. Central Point Software Inc., 15220 N.W. GreenBrier Pkwy., #200, Beaverton, Or. 97006; (503) 690-8090. Price: $149.

Hardware requirements: IBM PC, XT, AT, PS2 or 100% compatible with DOS 3.0 or higher and 512KB RAM. Hard disk recommended. Supports Microsoft mouse driver version 6.14 or higher, Logitech/Dexxa mouse drivers version 3.4X or higher, or 100% compatible.

There is something comforting about a Swiss Army Knife. In a single package you can find a variety of tools that might prove useful in many situations. Though you may never use some of them, there is comfort in knowing that the tools are there if and when they are needed. A Swiss Army Knife is also quite compact. While each tool in the knife might have a more capable version somewhere at home, the fact that the knife bundles so much solid utility together in a small package is one of its major selling points.

So it is with PC Tools Deluxe, version 6. Central Point Software has managed to pack a large number of utilities densely into a single, concise package. Any given utility, like its database manager, most certainly has larger and more exhaustive cousins such as dBASE, Paradox, DataEase, or Reflex. However, there is perhaps no other single program today that squeezes so many functions into such a small package.

There have been several, exhaustive reviews of PC Tools 6. This review does not duplicate that effort (see the subheading titled "Other Reviews" and the end of this article). Instead, I will give a general analysis of the program's major areas and list a few of its pluses and minuses.

PC Tools is organized around three basic sets of functions: a DOS shell, a set of extra utilities, and a desktop organizer. The desktop and the DOS shell can be run as terminate-and-stay-resident (TSR) programs. This means they can be loaded into RAM and thus are available while most other programs are running.

DOS Shell

This subsystem of PC Tools is called PC Shell. It is used primarily to facilitate the maintenance and organization of hard disks. As with other DOS shells, PC Shell attempts to eliminate the necessity of learning the DOS command structure. This is accomplished through the display of a disk's directory tree structure and file names. Operations are chosen by pressing hot keys or using "point-and-shoot," pull-down menus. PC Shell offers three different levels of user interface: beginner, intermediate, and advanced. A collection of native format data file viewers is extremely useful. An extra bonus for laptop users is a copy of LapLink, which uses the PC Tools interface and a cable to transfer information between the laptop and a desktop computer. The DOS Shell can be run in resident mode.

Extra Utilities

The set of extra utilities in PC Tools is not organized strictly as such; rather all of them are capable of being run as separate programs. As with PC Shell, most of them are concerned with the maintenance and organization of the hard disk. Compress is a program that optimizes your hard disk for faster performance by reducing file fragmentation. PC Backup is a full-featured backup utility. PC Format is used for formatting floppy disks. PC Secure will encrypt and decrypt files according to U.S. Department of Defense (DES) specifications. Diskfix is a utility that will read, analyze, make suggestions, and fix many errors or other likely trouble spots on the floppy disks. Undelete will allow you to recover files that were mistakenly deleted. Mirror can aid in this process by keeping a record of those files that have been deleted. PC Cache can speed up the reading and writing of the hard disk by letting the user set up an area in RAM to store frequently-accessed data.

Desktop Utilities

The desktop utilities in PC Tools are part of the Desktop Manager. These are useful programs that are similar to more expensive, stand-alone programs. As such, they are sort of mini versions of other software. Notepad is a mini word processor that allows one to draft, spell check, and print documents. Outliner is simply a notepad that allows the user to organize thoughts through the expansion, collapse, promotion, and deletion of different headings. Database is a mini database manager that can be used to create, sort, delete from, and add to lists of data. The data are stored in dBASE-compatible .DBF format. Telecommunications will allows a user to make connections to other computers through phone lines. There are four calculators available. One is a simple algebraic calculator, which is similar to the simple models found for under ten dollars in most stores. The other three are based on the sophisticated Hewlett Packard series: the HP-12C financial calculator, the HP-11C scientific calculator, and the HP-16C programmers calculator. Though none of these calculators implements all functions with the ease of the original HPs, they are still useful for many tasks, whether the need is for calculating the yield-to-maturity of a bond, doing polar conversions, or 2's complement arithmetic. The appointment scheduler is helpful is showing a five-day schedule at a glance, alerting the user to conflicts, finding free time, and allowing the user to attach lengthy notes to any appointment. A clipboard cuts and pastes both among the PC Tools and among other programs. Finally, the Macro Editor creates simple ASCII text files that instruct PC Tools to play specified keystrokes whenever some hot-key combination is pressed. This allows you to run complicated sets of PC Tools operations easily. If the Desktop Manager is running in resident mode, you can run these macros from almost any application.

The Bright Spots

There are many shining features of PC Tools. The following lists some of its best attributes.

PC Backup: Though PC Backup is only one of a number of utilities packaged with PC

Tools, it has received superior ratings in independent software reviews when compared to other, stand-alone backup utilities. It is one of the strongest of Central Point's offerings. A manual that is specific to the backup utility is included. PC Backup is so helpful that it alone is worth the price of PC Tools. (Note: a new stand-alone version of PC Backup is now available.)

File Viewers: PC Tools comes with thirty-five data file viewers. A data file viewer shows data from a specific application in that application's native format. For example, when the DOS type command is used to look at the data file for a Lotus spreadsheet (extension .WKS or .WK1), you see a lot of garbage in addition to the needed data. This is binary data that is part of the formatting instructions for the file. The PC Tools' Lotus 1-2-3 file viewer, however, allows you to look at the file in the familiar Lotus style. It is even possible to move the cursor around to see cells that are not on the screen. For example, if you are working in Microsoft Word and need to take a quick look at a Quattro spreadsheet, the file viewers make this task simple. If the PC Shell is running resident mode, it is much easier to use the Quattro data file viewer than to quit Word (remembering to save the document), start Quattro, pull up the spreadsheet, look at your data, close Quattro, start Word again, and retrieve your document. The advantage of a data file viewer becomes evident the longer the hard disk is used. As more and more files are stored on the disk, the file viewers become an important disk management tool, allowing you to see what is behind that file name before it is moved, copied, renamed, or deleted. There are seven database viewers, nine spreadsheet viewers, eleven word processor viewers, six compression viewers, a binary file viewer, and a .PCX graphics file viewer.

The Tarnish

Perhaps the biggest disappointments in using PC Tools are a result of the law of increasing expectations. With all that it does, you might be surprised if you don't find features that are available in your favorite word processor or DBMS. PC Tools stresses breadth and conciseness over depth and comprehensiveness. There are, however, a few small areas to watch out for.

Database: Plan well when you create a new database structure, because PC Tools will not allow you to restructure your file. If you decide that you would like to add a field to your existing name and address database, find someone with a copy of dBASE.

Notepad: When you create a document in Notepad with the word wrap on and save the file in ASCII text mode, no hard returns are inserted at the end of lines. This can be a problem when using Notepad to draft text files off-line for uploading to a host utility such as e-mail at a computing center or bulletin board.

PC Backup: It is very easy to teach a user how to employ the point-and-shoot method of selecting files and directories that need backing up. However, PC Backup will not allow you to save that set of files for future backups. In order to save a backup set, you must list file specifications through the "Include/Exclude" menu choice. This is harder to teach to users unfamiliar with DOS path names.

Documentation and Support

The documentation for PC Tools is quite good including online help while running the program, and three well-written manuals, namely one for PC Shell, one for the Desktop Manager, and one for PC Backup.

Central Point Software's technical support department can be reached via mail, FAX, or phone (8 a.m.–5 p.m., Monday–Friday). Their technical support staff is generally well-informed and helpful. If a technician is unavailable within the first five minutes, the user can leave a message, and a staff member will call back usually within fifteen minutes. The Support Department also runs a bulletin board containing tips, updates, and new information.

Summary

In conclusion, PC Tools Deluxe version 6 has come to dominate the DOS Shell/Utilities/Backup market for a good reason. Central Point Software has bundled an astounding number of useful features into a well-integrated, low-cost package. It might cost you five or six times more to purchase all the software that would do what PC Tools can do. This makes it a must for users of all DOS machines, particularly those with a small budget for software.

Other Reviews

Rarely does a software package receive rave reviews across the board. PC Tools Deluxe version 6 is in that enviable position. (With utility software predicted to be a real growth area, look for Peter Norton Computing, now under the Symantec wing, to try to take some of Central Point's substantial market share.) Listed below are more reviews of PC Tools Deluxe version 6.

Bigley, T. "Backup for Safe Keeping." *InfoWorld* 12, no.42:67–81 (Oct. 15, 1990).

———."PC Tools Deluxe Pushes Utility Features Envelope." *InfoWorld* 12, no.19:66–67 (May 7, 1990).

Brown, B. "Playing the DOS Shell Game." *PC Magazine* 9, no.11:185–240 (June 12, 1990).

Marshall, P. "Giving DOS a New Face." *InfoWorld* 12, no.16:57–73 (Apr. 15, 1990).

Specter, L. "Making DOS Manageable." *PC World* 8, no.7:119–30 (July 1990).

Walkenbach, J. "No Excuses Backup Software." *PC World* 8, no.7:149–63 (July 1990).—*Rick Gates, University of Oregon.* ■■

Other Recent Receipts

Listed here are books and other publications received for review that are of potential interest to LITA members. Some of these materials may be reviewed in later issues of ITAL.

Academic Library Management. Ed. by Maurice B. Line. Chicago: American Library Assn., 1990. 216p. $50 (ISBN 0-85365-839-0). ALA Order Code L839-0-0010.

Aluri, Rao, and Donald E. Riggs. *Expert Systems in Libraries.* Norwood, N.J.: Ablex, 1990. 342p. (ISBN 0-89391-589-0).

Annual Review of Information Science and Technology. V.25. Ed. by Martha E. Williams. New York: Elsevier, 1990. 491p. $104.50 (ISBN 0-444-88531-5).

Beyond the Book: Extending MARC for Subject Access. Ed. by Toni Petersen and Pat Molholt. Boston: Hall, 1990. 230p. paper, $39.95 (ISBN 0-8161-1924-4).

Boss, Richard W. *Library Manager's Guide to Automation, 3d ed.* Boston: Hall, 1990. 202p. paper, $24.95. (ISBN 0-8161-1943-0).

Case Studies of Optical Storage Applications. Ed. by Judith Paris Roth. Westport, Conn.: Meckler, 1990. 139p. $45 (ISBN 0-88736-535-3).

CD-ROM Licensing and Copyright Issues for Libraries. Supplement to Computers in Libraries, no.23. Ed. by Meta Nissley and Nancy Melin Nelson. Westport, Conn.: Meckler, 1990. 95p. $34.95 (ISBN 0-88736-701-1).

Converting Information for W.O.R.M. Optical Storage: A Case Study Approach. Ed. by Judith Paris Roth. Westport, Conn.: Meckler, 1990. 284p. $49.50 (ISBN 0-88736-380-6).

Dewey, Patrick. *Public Access Microcomputers: A Handbook for Librarians.* 2d ed. Boston: Hall, 1990. 155p. paper, $35 (ISBN 0-81611-896-5).

Directory of Library and Information Retrieval Software for Microcomputers. 4th ed. Comp. by Hilary Dyer and Alison Gunson. Brookfield, Vt.: Gower, 1990. 134p. paper, $38.95 (ISBN 0-566-03628-2).

DISC Magazine (Monthly, Jan. 1991–) Ed. by Linda Helgerson. Falls Church, Va.: Helgerson Associates, Inc. (510 N. Washington St., Suite 401, Falls Church, VA 22046-3537; (703) 237-0682), 1990. $44.95/yr. (ISSN 1052-4053).

Eaton, Nancy, Linda MacDonald, and Mara R. Saule. *CD-ROM and Other Optical Information Systems: Implementation Issues for Librarians.* Phoenix, Ariz.: Oryx, 1989. 168p. paper, $32.50 (ISBN 0-89774-448-9).

European Library Networks. Ed. by Karl W. Neubauer and Ester R. Dyer. Norwood, N.J.: Ablex, 1990. 448p. $75 institutional; $42.50 personal (ISBN 0-89391-157-7).

Fox, David. *The CD-ROM Market in Canadian Libraries.* Westport, Conn: Meckler, 1990. 61p. paper, $19.95 (ISBN 0-88736-734-8).

The Future of Data Dictionaries. Ed. by Simon Halloway. Brookfield, Vt.: Gower, 1990. 121p. paper, $39.95 (ISBN 0-566-09020-1).

Getting in Print, Staying in Print. Ed. by William A. Forster. HERTIS (Hatfield Polytechnic Library, College Lane, Hatfield, Hertfordshire AL10 9AD), 1990. 112p. paper, £12.50 (ISBN 0-85267-284-5).

Greenia, Mark W. *Computer Abbreviations and Acronyms: A Guide for Managers, Consultants, Students and Computer Professionals.* Sacramento, Calif.: Lexikon Services (8042 Singletary Way, Sacramento, CA), 1990. 97p. paper, $20 (ISBN 0-944601-14-6).

Greenia, Mark W. *Computer Security Information Sourcebook: A Guide for Managers, Attorneys, and Other Concerned Professionals.* Sacramento, Calif.: Lexikon Services (8042 Singletary Way, Sacramento, CA), 1990. 55p. paper, $20 (ISBN 0-944601-16-2).

Information Ethics: Concerns for Librarianship and the Information Industry. Ed. by. Anne P. Mintz. Jefferson, N.C.: McFarland, 1990. 96p. paper, $11.95 (ISBN 0-89950-514-7).

Lane, Elizabeth S. *Microcomputer Management & Maintenance for Libraries.* Westport, Conn.: Meckler, 1990. 205p. $39.50 (ISBN 0-88736-522-1).

The Librarian's Guide to Wordperfect 5.0. Ed. by Cynthia LaPier. Westport, Conn.: Meckler, 1990. 177p. $39.50 (ISBN 0-88736-4).

The Library Microcomputer Environment: Management Issues. Ed. by Sheila Intner and Jane Anne Hannigan. Phoenix, Ariz.: Oryx, 1988. 272p. paper, $29.95 (ISBN 0-897740229-X).

Library Perspectives on NREN (The National Research and Education Network). Ed. by Carol A. Parkhurst. Chicago: Library and Information Technology Assn., 1990. 75p. paper, $10.50 (ISBN 0-8389-7477-5).

Lopez, Victor, and Kenneth Ansley. *Free and User Supported Software for the IBM PC: A Resource Guide for Libraries and Individuals.* Jefferson, N.C.: McFarland (Box 611, Jefferson, NC 28640; (919) 246-4460), 1990. 208p. paper, $20.95 (ISBN 0-89950-499-X).

Mueller, Michael, and Martin Schuchardt. *The Leisure Suit Larry Story.* Grand Rapids, Mich.: Abacus, 1990. 156p. paper, $14.95 (ISBN 1-55755-086-7).

Network Modeling, Simulation, and Analysis. Ed. by Ricard F. Farzia and Mario R. Garzia. Electrical Engineering and Electronics Series, no.61. New York: Marcel Dekker, 1990. 392p. $99.74 U.S. & Canada; $119.50 elsewhere (ISBN 0-8247-7876-6).

Nicholls, Paul T. *CD-ROM Collection Builder's Toolkit: A Complete Handbook of Tools for Evaluating CD-ROMs.* Weston, Conn.: Pemberton Pr., 1990. 180p. paper, $29.95 (ISBN 910965-01-3).

Pettersson, Rune. *Visuals for Information.* Englewood Cliffs, N.J.: Educational Technology Publications (720 Palisade Ave., Englewood Cliffs, N.J. 07632), 1989. 315p. $37.95 (ISBN 0-87778-218-0).

Saffady, William. *Optical Disks vs. Magnetic Storage.* Westport, Conn.: Meckler, 1990. 122p. paper, $42.50 (ISBN 0-88736-703-8).

Tenopir, Carol, and Jung Soon Ro. *Full Text Databases.* New Directions in Information Management, no.21. Westport, Conn.: Greenwood, 1990. 268p. $39.95 (ISBN 0-313-26303-5). ■■

INSTRUCTIONS TO AUTHORS

Information Technology and Libraries welcomes manuscripts related to all aspects of library information technology. Some specific topics of interest are mentioned on the masthead page. Feature articles, communications, letters to the editor, and news items are all considered for inclusion in the journal. Feature articles are refereed; other items generally are not. All material is edited as necessary for clarity and length.

Manuscripts must be typewritten and the original submitted with one duplicate. All text must be double-spaced, including *footnotes and references.* Manuscripts should conform to *The Chicago Manual of Style,* 13th ed., rev. (Chicago: Univ. of Chicago Pr., 1982). Illustrations should be prepared carefully as camera-ready copy, neatly drawn in a professional manner on separate sheets of paper. Manuscript pages, bibliographic references, tables, and figures should all be numbered consecutively. Final versions of papers accepted for publication must be submitted on floppy disk (5.25" or 3.5") in *WordPerfect* (DOS versions 4.42, 5.0, and 5.1 accepted) along with two printed copies. If *WordPerfect* is not available, an ASCII copy is acceptable.

Feature Articles consist of original research, state-of-the-art reviews, or comprehensive and in-depth analyses. An abstract of one hundred words or less should accompany the article on a separate sheet. Headings should be used to identify major sections. Authors are encouraged to relate their work to other research in the field and to the larger context of economic, organizational, or management issues surrounding the development, implementation, and use of particular technologies.

Communications consist of brief research reports, technical findings, and application notes. An abstract need not be included.

Letters to the Editor may offer corrections, clarifications, and additions to previously published material or may be independent expressions of opinion or fact related to current matters of concern in the interest area of the journal. A letter commenting on an article in the journal is shared with the author, and a response from the author may appear with the letter.

News and Announcement items may announce publications, conferences, meetings, products, services, or other items of note.

Book Reviews are assigned by the book review editor. Readers wishing to review books for the journal are invited to contact the book review editor, indicating their special areas of interest and expertise.

Software Reviews are assigned by the software review editor. Readers wishing to review library-related software for the journal are invited to contact the software review editor, indicating their special areas of interest and expertise.

Names and addresses of the journal editors may be found in paragraph three on the masthead page. In all correspondence please include your own name, institutional affiliation, mailing address, and phone number.