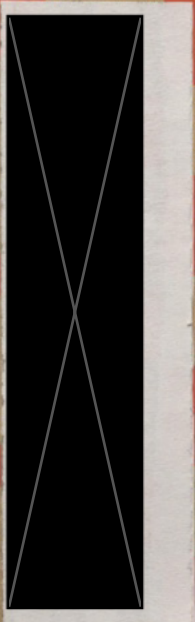


Information Technology and Libraries

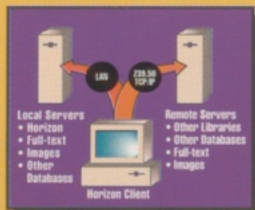
Volume 14 Number 3 ISSN 0730-9295

September 1995



The standard by which other client/server systems are judged

Horizon™ is the information management system that sets the standard by giving your library all the benefits of client/server computing.



A true client/server system

The flexibility of an open systems environment means you can add to the system as needs change and adapt to future technologies while leveraging your investment.

Expanded patron access

In addition to your own Horizon database, patrons can use Horizon to gather bibliographic data from any Z39.50-compliant system and connect to global resources over the Internet.

Graphical user interfaces

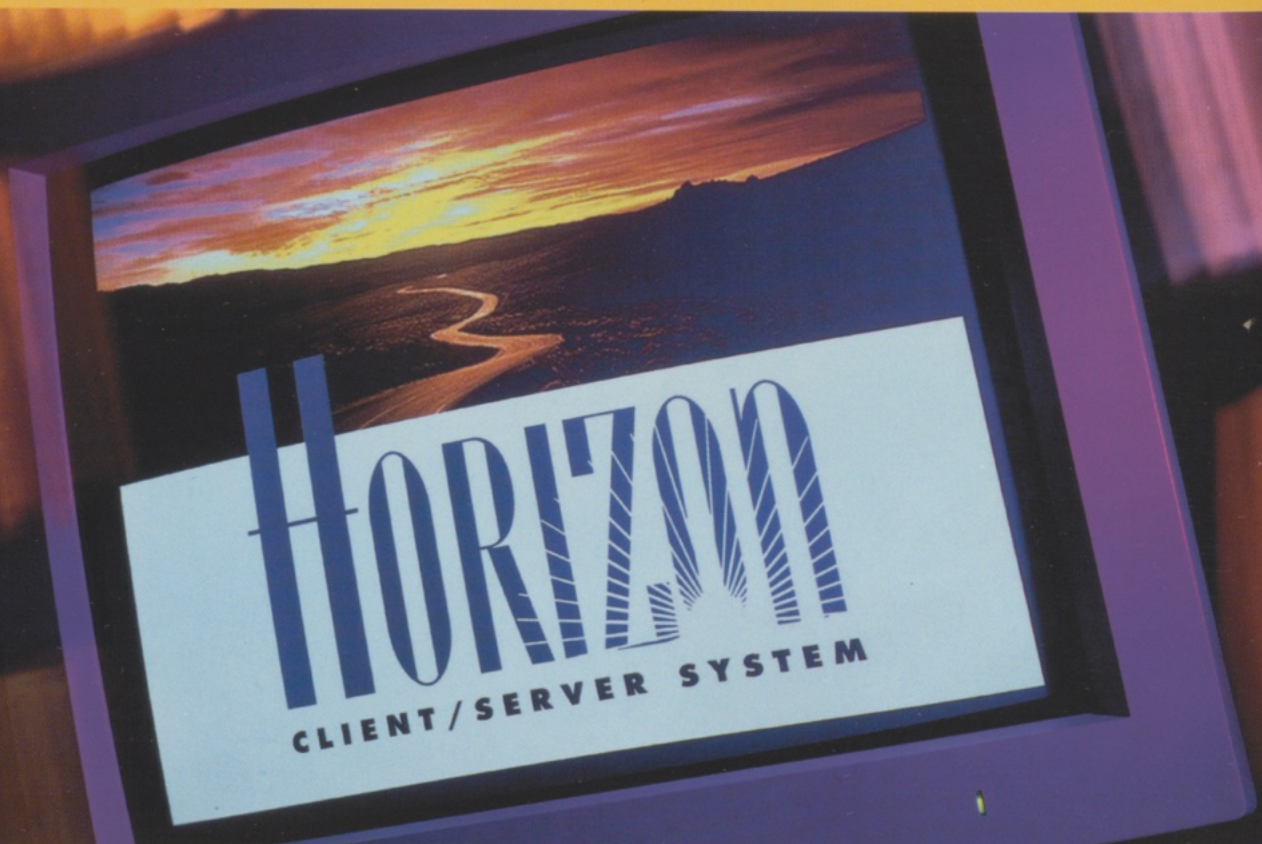
The Horizon client software offers the familiarity of Microsoft Windows®, System 7 (Macintosh®), or OS/2® interfaces. Patrons can search effortlessly with icons, pull-down menus, and point-and-click commands.

Proven by libraries like yours

You need a system with a history of success, one designed by library professionals. Horizon has met those requirements with thorough testing by libraries around the world.

Horizon is available today

You can bring the benefits of Horizon to your library today. Let us show you the leading client/server system.




Ameritech Library Services

Academic Division
1007 Church Street
Evanston, IL 60201-3665
(800) 556-6847
Fax (708) 866-0178
academic@amlibs.com

Special & Public Divisions
400 Dynix Drive
Provo, UT 84604-5650
(800) 288-8020
Fax (801) 223-5202
special@amlibs.com
public@amlibs.com

Ameritech

If the Holy Grail was in your  library, would it take another

two thousand  years to put your finger on it? Not if you were

using the GALAXY® system. You'd find it fast, along with related



literature, Internet



-discussion groups and

local meetings of Grail enthusiasts. Because GALAXY has state-of-

the-art



modules for Online Public



Access

Catalog, journal citation retrieval, networking, community informa-

tion, circulation control, acquisitions, serials, cataloging and much

more. All integrated for maximum performance and ease of use.

To speed all your quests, call 1-800-962-9580.

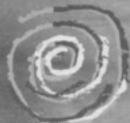
GALAXY®

Solutions for Today's Libraries

GAYLORD
INFORMATION
SYSTEMS

Gaylord

The Trusted Source



Collection Management for the 90's and beyond

BLACKWELL'S
NT|Online

Blackwell's **New Titles Online** database includes **Tables of Contents** and **Descriptive Summaries** for new and forthcoming scholarly monographs.

- Browse the contents pages before placing your order
- Examine the publishers title description
- Order the title electronically
- Insure title coverage on approval
- Search both Blackwell North America and B.H. Blackwell New Title databases
- View full MARC Records
- Access through the Internet

FOR A FREE TRIAL OF THE WORLD'S MOST COMPREHENSIVE
ACADEMIC COLLECTION MANAGEMENT SYSTEM,
CONTACT YOUR BLACKWELL REPRESENTATIVE.

BLACKWELL
NORTH AMERICA, INC.

6024 SW Jean Road, Building G, Lake Oswego, Oregon 97035.

Telephone: 800-547-6426 Fax: 503-639-2481

E-mail address: BRIDGES@BNAME.BLACKWELL.COM

TECHNOLOGY DEDICATED TO THE LIBRARIES OF TOMORROW

Thomas W. Leonhardt, Editor

University of Oklahoma, Norman

Marjorie E. Bloss, Managing Editor

Center for Research Libraries, Chicago, Illinois

Susan B. Harrison, Book Review Editor

New York Public Library, New York

George S. Machovec, Software Review Editor

CARL, Denver, Colorado

EDITORIAL BOARD: Joan M. Aliprand, Research Libraries Group, Mountain View, California; Walt Crawford, Research Libraries Group, Mountain View, California; Patrick Flannery, Houston Academy of Medicine, Texas Medical Center Library, Houston; Charles Hildreth, University of Oklahoma, Norman; Gail Junion-Metz, unaffiliated, ex officio; Katherine Klemperer, Harvard University; Ben-Ami Lipetz, School of Information Science and Policy, State University of New York, Albany; Marilyn Lutz, University of Maine, Orono; Randy L. Pedersen, University of North Dakota, Grand Forks; V. Louise Saylor, Eastern Washington State University, Cheney; Xiao-Yan Shen, Santa Clara University, California; John Webb, Washington State University, Pullman.

Information Technology and Libraries (ISSN 0730-9295) 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 70 of the March 1995 issue. Manuscripts of articles, communications, and news items should be addressed to the editor-designate: James J. Kopp, Library Director, Wilson W. Clark Memorial Library, University of Portland, P.O. Box 83017, 5000 N. Willamette Blvd., Portland, OR 97283-0017. 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, Colorado Association of Research Libraries, 3801 E. Florida Ave., Ste. 370, Denver, CO 80210-2542. ADVERTISING: Contact Todd Goldman, Todd Goldman Associates, 3418 Handy Road, Suite 201, Tampa, FL 33618; (813) 264-2772.

As a matter of policy, *Information Technology and Libraries*, as the scholarly organ of LITA, does not review LITA publications. Notice of new publications from LITA will generally be found in the "Other Recent Receipts" column following reviews.

Information Technology and Libraries is a requisite of membership in the Library and Information Technology Association. Subscription price, \$22.50, is included in membership dues. Nonmembers may subscribe for \$50 per year in the U.S.; \$55 in Canada, Mexico, Spain, and other PUAS countries; \$60 in other foreign countries. Single copies, \$15. 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 Production Services (Eileen Mahoney, Bruce Frausto, and Donovan Vicha),

Information Technology and Libraries

Volume 14, Number 3

September 1995

ISSN 0730-9295

- 132 Editorial
THOMAS W. LEONHARDT
- 135 Object-Oriented Cataloging
MICHAEL HEANEY
- 155 What Will Collection Developers Do?
MICHAEL BUCKLAND
- 161 Precision and Recall in Title Keyword Searches
MONICA CAHILL MCJUNKIN
- 173 COMMUNICATIONS
- 173 "Two Thumbs Up!" Library Applications for Video Technology
DAREN CALLAHAN
- 176 Planning for Sustainable Automation
BRUCE HARLEY
- 179 OPAC Database Creation Problems
NADINE L. BAER, JAMES A. BARRETT, AND KARL E. JOHNSON
- 185 Migration from LS/2000 to KeyNOTIS
JOYCE L. HUANG AND TING ZHENG
- 190 Internet Use through the University of Toronto Library
JOY TILLOTSON, JOAN CHERRY, AND MARSHALL CLINTON
- 199 Letters and Announcement
- 201 Book Reviews
- 198 Index to Advertisers

Cover Design by Jim Lange. Interior Design by Dianne M. Rooney

American Library Association, 50 E. Huron St., Chicago, IL 60611.

Publication of material in *Information Technology and Libraries* does not constitute official endorsement by LITA or the ALA.

Abstracted in *Computer & Information Systems, Computing Reviews, Information Science Abstracts, Library & Information Science Abstracts, Referativnyi Zhurnal, Nauchnaya i Tekhnicheskaya Informatsiya, Otdyelnyy 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 avail-

able 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 Z39.48-1992. ∞

Copyright © 1995 American Library Association. All material in this journal subject to copyright by ALA may be photocopied for the noncommercial purpose of scientific or educational advancement granted by Sections 107 and 108 of the Copyright Revision Act of 1976. For other reprinting, photocopying, or translating, address requests to the ALA Office of Rights and Permissions.

Poor Johnny One-Note: Speaking Up for Libraries

Two lives ago, I had the wonderful experience of playing in the 84th U.S. Army Band, Fulda, Germany. Technically, I was a professional musician for a year, but I never forgot my strictly amateur talent and my good fortune to be sitting in with some true professionals.

One of my friends in the band was a reed man—tenor sax and clarinet. He had read or heard somewhere that John Coltrane would play an individual note over and over again until he felt he had captured its essence. My friend, emulating his idol, would play a single note or three-note riff for what seemed like hours at a time until he and everyone around him knew the essence of what he played. He tried it on our piano, too, until someone hid the key from him.

For the past year or so in writings and speeches, I have felt as if I have been poor Johnny One-note, seeking the essence of libraries by playing my three-note riff—books, libraries, librarians. I was slowly discovering the essence of this riff, but did anyone in the audience dig it? This spring I was reassured and gratified when I received a letter from an academic librarian in Ohio who had heard my riff at a conference last fall. The piece that I played stressed the importance of libraries, imploring all within earshot to evaluate technology and the Internet critically and not simply accept them as the salvation of librarians who want to ply their trade in the future.

As I was composing this editorial, President Clinton was imploring us to speak out against the purveyors of hate, intolerance, and bigotry. In a free society we are all responsible for our freedom, and to remain silent when others misspeak in our names is to tacitly approve their messages.

Librarians cannot remain silent either when the importance of libraries, reading, and free and equal access to knowledge is dismissed as outmoded and old-fashioned. And we cannot remain silent when librarians who insist that “librarian” is an honorable title are said to have their traditionalist heads buried in the sand.

Librarians are finally beginning to speak up and others are hearing the ring of truth. First we had Roma Harris and *Librarianship: the Erosion of a Woman's Profession*,¹ followed shortly by *The Myth of the Electronic Library: Librarianship and Social Change in America*, by William Birdsall.² Now we have *Future Libraries: Dreams, Madness, and Reality* by Walt Crawford and Michael Gorman.³

Gorman and Crawford have formulated Five New Laws of Library Science, “a reinterpretation of Ranganathan’s [Five Laws of Library Science] in the context of the library of today and its likely futures.”

Libraries serve humanity.

Respect *all* forms by which knowledge is communicated.

Use technology intelligently to enhance service.

Protect free access to knowledge.

Honor the past and create the future.⁴

The silence has been broken. It is time to take a stand for libraries, books, reading—and librarians.

Technology and the Internet are already playing important roles in libraries, and librarians are using these new tools creatively to expand their services. But as Crawford and Gorman tell us, “It is human to indulge in hyperbole and oversimplification at times. It is foolishness to do it all the time.”⁵ When it comes to the Internet and the so-called virtual library, it is time to stop our foolishness and remember who we are and who we serve.

A well-informed society is the surest way to preserve our freedom. In *Prescribing the Life of the Mind*, Charles W. Anderson suggests that education is wasted on the young and that it ought to be available to all of us as we go through life. “Perhaps there should be universities on every street corner, as there now are churches, and it would be assumed that many would attend them, as they now do churches, regularly throughout life.”⁶ There are such places, Professor Anderson; they are called public libraries and, yes, many attend them regularly throughout their lives.

Notes

1. Roma M. Harris, *Librarianship: The Erosion of a Woman's Profession* (Norwood, N.J.: Ablex, 1992).

2. William F. Birdsall, *The Myth of the Electronic Library: Librarianship and Social Change in America* (Westport, Conn.: Greenwood, 1994).

3. Walt Crawford and Michael Gorman, *Future Libraries: Dreams, Madness, and Reality* (Chicago: ALA, 1995).

4. *Ibid.*, 7–8.

5. *Ibid.*, 40.

6. Charles W. Anderson, *Prescribing the Life of the Mind: An Essay on the Purpose of the University, the Aims of Liberal Education, the Competence of Citizens, and the Cultivation of Practical Reason* (Madison, Wis.: University of Wisconsin Press, 1993), 94.



INFORMATION SOLUTIONS



Looking for **CD-ROM** solutions?

Our library-oriented staff has the experience and expertise to guide you in selecting the most effective CD-ROM system for your library.

HITACHI

CDR-6700/6750 internal
CDR-1900S/1950S external
2 year manufacturer's warranty

TODD

TCDR-2XM/2XMS external
4 to 8 drive modular towers
Cabinets for up to 64 drives
3 year manufacturer's warranty

PIONEER

DRM-602X 6 disc changer
DRM-604X 6 disc changer
DRM-1804X 18 disc changer
1 year manufacturer's warranty

10332 Old Olive Street, Saint Louis, Missouri 63141
Phone 800-264-3799 Fax 314-567-3798 Internet: AXXESS@MLNC.COM

* A subsidiary of MLNC *

THE ONLINE 100...

ONLINE Magazine's Field Guide to the 100 Most Important Online Databases

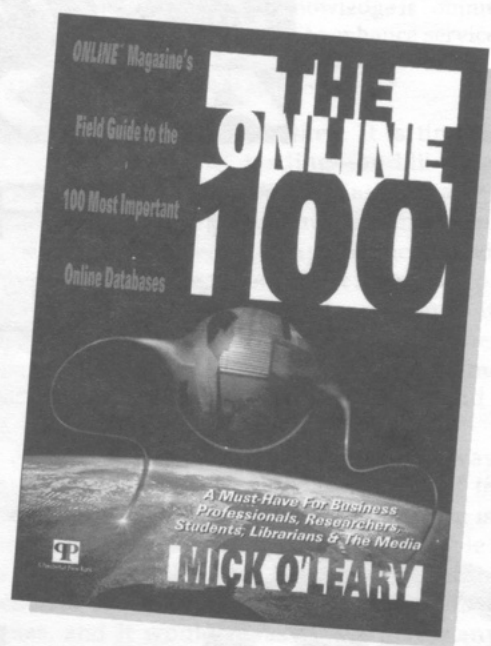
Absolutely anyone doing online research today needs **THE ONLINE 100**—A powerful and practical resource that unlocks the secrets of critical information spanning all areas of industry, education, science, and the humanities.

Author Mick O'Leary, a leading information industry expert, evaluates the 100 most important and useful databases and provides time and money-saving tips including:

- concise explanations of each database's strengths and weaknesses
- typical search costs
- critical information on span and update frequency of data
- hard-to-find information on CD-ROM and Internet availability
- advice on when *not* to use a certain database—with pointers on where to look instead
- complete contact information for all principal databanks
- practical tips and advice to help you get the most out of a database

This is a brave and needed effort. It used to be that good searchers could know every database that was available, but this isn't possible anymore...I like the idea of the book, and the structure. THE ONLINE 100 will be very helpful to experienced searchers as well as to beginners.

—Carol Tenopir, Prof. of Information Sciences,
U. of Tennessee School of Information Sciences,
and Columnist, *Library Journal*



240 pages, \$22.95

Available Now from the Publisher
of ONLINE Magazine

To order or for a free
Pemberton Press Books catalog
call 1/800/248-8466
(outside U.S. call 203/761-1466)



Books & Tools for the Information Age

Catalogs have developed from lists of physical items present in particular libraries into computerized access and retrieval tools for works dispersed across local and national boundaries. Works themselves are no longer necessarily constrained in one physical form. Cataloging rules, however, have not evolved in parallel with these developments. This paper reanalyzes the nature of works and their publication in an approach based on object-oriented modeling, and demonstrates the advantages to be gained thereby. It suggests a strategic plan to enable an organic transformation to be made from current MARC-based cataloging to object-oriented cataloging.

Finding Books in Libraries

Catalogs began as listings of particular books on specific shelves. My own institution, the Bodleian Library, compiled the first such published catalog in the English-speaking world nearly four centuries ago (Bodleian Library 1986). Over the centuries since then—but slowly—ideas such as relative location, consistent description, and catalog headings (i.e., access points) have been developed. This process came to full flower with the publication of the first edition of the *Anglo-American Cataloguing Rules* in 1967. In the following year, MARC was initiated. Here were a complete set of cataloging rules and a comprehensive way of coding catalog entries for computer use, arriving more or less one right after the other. The revision of the rules for the second edition in 1978 and its official adoption in 1980 by the two major sources of bibliographic records in the United Kingdom and the United States spread that florescence yet further. As a result, AACR2 has taken firm root all over the Anglophone world and beyond.

The revision of the second edition of AACR in 1988 was not felt to be radical enough to warrant its identification as a full new edition. It really does seem that with the 1988 revision of AACR2 we have reached a plateau of maturity, with comprehensive rules based on firm theoretical foundations.

In reality, of course, we have reached no such plateau. The problems arise from the fact that library catalogs do not function solely as descriptive lists of books: they function also as elements in library management systems and, in the electronic age, as sophisticated information retrieval tools.

The catalog as a list of books is centered on the physical objects contained in a library. The structure of AACR2 makes it very clear that description is the first priority and that, once the description of an object is complete, one's attention may be turned to attaching "handles" to it so that readers may gain access to the material. The handles—the access points—arise from the description. The description itself is a stage on the road leading to the full text.

In some environments in which AACR2 is used, different aspects of it may be considered important. In most booksellers' catalogs, there is only one access point for each work, and the bibliographic description is the main element; after all, attention is being focused on the physical item that is being offered for sale. In a company research library, classified access may be more important than access by personal author. How valid is AACR2 for all of these situations? Is the code really equally suitable for all environments?

Access based on bibliographic description is historically grounded in the development of library catalogs. Until this century each library was self-contained and each catalog truly reflected the library's contents. If a requirement could not be met within the library, that was the end of the matter. The user had to be satisfied with the library's contents. But how many users enter a library looking for particular physical items? In the case of subject searches, the answer is obviously "None." Even in the case of what are generally called "known item" searches, most users are looking for any item which satisfies the request "an edition of Hamlet," rather than "the Macmillan edition of Hamlet."

The history of the development of the theoretical bases of cataloging has been rehearsed many times. Many of these threads have been drawn together by Eric Wainwright of Australia in what should be a seminal article for consideration of the role of cataloging in the online age (Wainwright 1991). He avers that "cataloguing has been based on the needs of individual libraries . . . i.e., has been based principally on the physical object in hand, as received by the library," that "early catalogues described relatively small collections . . . in a very limited number of manifestations," and that "the structure of the catalogue was prescribed by the technological limitations of manually prepared entries" (p. 8). He argues that none of these considerations is valid today and concludes by quoting Boll: "Catalogues and their rules should be for retrieval not description."

Michael Heaney is Head of Foreign-Language Cataloging, Bodleian Library, Oxford.

Development of MARC Formats and Authority Files

While a few national institutions distributed catalog cards to client libraries before the advent of MARC, it was the development of MARC that led to the practice of record acquisition and sharing on a wide scale. The MARC formats allow AACR2 (and other) data to be encoded, but they also allow coding of other elements: subject strings, coded information specifically for machine manipulation, and in some implementations copy or holdings information and cross-references.

The development of MARC has not been marked by the signs of maturity that characterize AACR2. US MARC and UK MARC are different in many ways, and it is not possible to translate from one to the other without loss of data. US MARC is the direct successor of the Library of Congress's preexisting program of supplying catalog cards and reflects that origin. (UK MARC has a more direct correlation with AACR2.) This fact is a salutary reminder of the fact that MARC, as originally developed, was not devised for online public access to and interrogation of databases, but as a method of automating the production of catalog cards and printed bibliographies. AACR2 and MARC remain, at their core, means of producing lists of books and, as libraries implement them, lists of books in a particular place.

Other countries have also produced their national versions of MARC, and in an attempt to take steps to remedy this problem UNIMARC has been developed. So far this "Esperanto" of MARC seems to be faring better than the real Esperanto. It must be said that, of the three formats mentioned, it does incorporate the most logical analysis of computer-based cataloging. US MARC reflects its origin as a practical means of generating catalog cards; UK MARC incorporates AACR2 most closely; but UNIMARC began as a more abstract assessment of the needs of computer catalog records. It too, however, remains tied to ISBD (the International Standard for Bibliographic Description) and the Paris Principles.

As time has gone on it has been inevitable that the scope of MARC should broaden. Particularly in the United States, MARC has now broadened to include a holdings format, an authority files format, and—more distant from the core library world—a community information format. There are also proposals to extend the US MARC structure to the library task of tracking circulation (*Considerations* 1993). In the UK, extensions to the UK MARC format have been published for book trade use. These extensions all focus on the need to exchange information about bibliographical matters above and beyond the original basis in the production of catalog cards.

The first major extension of MARC was to the coding of authority files. Authority files are the key to many access points in catalogs. The Library of Congress (LC) files are the most widespread in use. Names, uniform titles, and name-title entries are listed in the Names Authority File. The relational structures of names and their variants are codified, as is the source of information. The British Library's Name Authority List is much simpler in structure than the LC's file. Only names are listed, not uniform titles or name-title combinations. Only "see" and "see also" references are listed, not any of the other types of references suggested by AACR2. The British Library and the LC are engaged in discussions on the merging of their respective name authority files. The British Library is also involved in plans for a transnational authority file in collaboration with European partners. This will inevitably raise more questions about syndetic structure and what constitutes identity. These and many other initiatives show clearly that the question of access is beginning to loom larger in the public mind. This shift of emphasis from description to access is in part driven by economic realities, which have encouraged a philosophy of libraries as facilitators of access to information rather than purely as holders of information.

In the United States, discussions have also begun on what lies at the heart of this article: the cataloging of "multiple versions," embodied in the existence of facsimile and original versions of a work; the discussions have not fully examined the broader ramifications of such an approach. Instead of having two distinct records for a work and its reproduction, or conflating them into a single record, the multiple versions proposals would have the relationship represented hierarchically. The proposal "eliminates the need to repeat the description of the original in the description of the reproduction" (American Library Association 1992, 1).

Library Computer Applications

Access and Description in the Online Age

The creation of online catalogs has led occasionally to proposals such as the one to discard the idea of main entry and to accord all name headings attached to a work equal status. A recent example is Bierbaum's article in the January 1994 issue of *American Libraries*. These ideas give primacy to online access over the niceties of linear access in printed catalogs. They question the relevance of AACR2 in the online age. In particular, the main entry problem questions the appropriateness of choosing to

emphasize one element derived from the bibliographic description over another. But the issue is more complicated than that and involves not just the idea of "main" entry but also the adequate identification, within the record, of related works. However, the related works problem itself arises from the construction of AACR2 and MARC in terms of the production of linear strings of text.

Even within AACR2 there are sleights of hand which disguise the centrality of the issue of description versus access. Theoretically, access points should arise from the description of the physical item, but in some instances they depend upon decisions about the "nature" of a work: for example, the text versus commentary aspect or the original author versus revising author in AACR2 rule 21.12B. AACR2 does give guidance which uses the layout of the physical object as a touchstone, but in the end the decision is based on the cataloger's conception of what the work "really" is.

Another important example is the case where a known author is not named in the work. Here the cataloger is instructed to give the known author as the main entry and to add a note identifying the person. Adding the note to the record is access determining description.

In online catalogs new problems arise from the interplay of AACR2 and the implementation of the functional translation of MARC records into computer files and databases. In many catalogs the fact that, say, Charles Dickens wrote *Oliver Twist* is reflected only in the fact that the name and the title occur in the same record. Names and titles are seen as essentially separate parts of the bibliographic record. Although name-title added entries come closer to reflecting the real situation, there is the anomaly that the two-field structure main-name-entry+uniform-title corresponds (awkwardly) to the single-field structure added-name-entry+uniform-title.

Another consequence of the lack of a name-title link is that, in many catalogs, if you perform a search asking for records with Dickens as author and *Oliver Twist* as title, you will retrieve a set of records displayed in a one-line display of name and title, in which all the names are Dickens, and all the titles *Oliver Twist*. Naturally, this is what was asked for; but if you know you will retrieve editions of *Oliver Twist*, should not the display of retrieved records presume that and concentrate on showing you the elements that will distinguish between the records?

The problems caused by the failure of MARC records to recognize the close structural link between name and title in a bibliographic record have frequently given rise to suggestions for change. The most recent has been MARBI discussion paper number 72, which proposed

that the 240 and 243 fields be abandoned in favor of subfields in the 1XX field (Possible 1994). This proposal was rejected on the grounds of the existence of long-standing and extensive implementations of the present configurations and the use of the 1XX in the bibliographic record, where it is intended to "head" all the 2XX titles when they are used.

Many systems allow users to place reservations on books. Usually systems allow the reservation to be placed either on the first copy of a given book to become available or on a specific copy. Most of these rely on identifying and flagging a given bibliographic record. How many systems allow users to reserve the first copy of any edition of a book which becomes available?

With the development of multimedia, it is now more and more common for the "same" work to be present in different manifestations: the book and the cassette recording, the CD and the video, or the book and the CD-ROM (as with the *Oxford English Dictionary*).

Many of the deficiencies of MARC in the online environment have been reviewed more than once. Michael Gorman recently proposed the development of "HYPERMARC" as one answer to the challenge that "if cataloguing rules . . . are to evolve to a higher plane, it is essential that the MARC format be reevaluated and radically recast" (Gorman 1992, 91).

Record Structure in Library Computer Applications

The majority of library systems developed in the 1970s and 1980s relied on inverted index files—information from records copied into indexes. This was the machine equivalent of the development of catalogs from shelflists. Some more forward-looking systems from the same period integrated the name fields (1XX, 7XX) in records with their name index, sometimes extending this approach to titles and subjects. These are steps towards a database organization.

Now that we are all interested in open systems, the most up-to-date library computer applications rely on widely available database management systems such as Oracle. Most are relational databases, some are extended relational databases. There are indications, however, that library catalogs are not handled particularly well by relational databases. The MARC record is immensely complicated. Relational databases work best with simple conceptual links between simple data elements. While such databases can, of course, incorporate all of the bibliographic elements and their links, they do so at the cost of requiring multiple database accesses to answer even straightforward bibliographical queries.

Object-Oriented Modeling

The Basis of Object-Oriented Modeling

This is where object-oriented databases and modeling come in. Object-oriented modeling bundles together in a single concept both the "entity" and what it can do: for example, the knowledge of what a person is, the characteristics of persons (e.g., that they have names), and the actions they can perform (e.g., they can write books). While the term is often used synonymously with "entity-relationship modeling," the difference is purely one of emphasis. Object-oriented modeling emphasizes the primacy of the object, or entity, over its relationships and is perhaps more abstract, not prescribing any particular method of implementation. It is a method of conceptual modeling not tied to any computer application, although many computer applications and programming tools now use its approach.

Object-oriented databases should be better than relational databases at handling complex entities such as books. They allow for inheritance of characteristics so that, for example, the characteristics of the entity "Oliver Twist" can be inherited by every publication that includes that entity. However, object-oriented databases are not yet well developed. There are as yet no agreed standards for them, and some technical problems remain—for example, there is no agreed calculus and algebra of objects. Some of the extended relational databases do work faster than has been the case with other types of databases, yet the way ahead seems to lie with hardware developments like massively parallel processing, rather than in matching database structures closely with conceptual models.

Object-oriented modeling emphasizes the identification of real-world entities and their relationships, which should remain stable even though their implementation may change. "Oliver Twist," for instance, is stable regardless of whether it manifests itself as a printed book, a microform, an electronic text, or an audio cassette. If the object, or entity, has been correctly analyzed, it should remain usable over a wide variety of changing circumstances and events, whether these affect the objects themselves or simply affect descriptions of them.

Object-Orientation and Library Catalogs

Object-oriented modeling seems in many respects ideal for the implementation of library catalogs, and these ideas are already beginning to find their way into general library automation packages. However, if the "object" is the MARC record based on AACR2, then either a

large amount of "deconstruction" of each record has to take place or the benefits of object-oriented modeling will be confined to nonbibliographical transactions such as circulation control. Some systems under development leave the MARC record intact (as a BLOB, a Binary Large Object), but copy all the relevant fields into object classes. Such a system obviously has cost overheads and massive redundancy in storage, and complicates the process of updating records.

I would argue that there are benefits to be had from a fundamental restructuring of MARC in order to allow it to reap the benefits of object-orientation. Even more, I would argue that there are benefits to be had from reinterpreting AACR2 from an object-oriented viewpoint, so as to ground it in considerations of access requirements and record sharing, rather than bibliographical description.

A 1989 article by Barbara Tillett represented early recognition of the potential of object-oriented cataloging, which she described under the term "entity-relationship model." However, Tillett used the model not to redefine the elements of the bibliographic world, but merely to describe the associations between name authorities, bibliographic records, and access and subject descriptors. She recognized the inappropriateness of, for example, treating variant titles as elements of bibliographical records and not as elements of an authority file. She wrote: "Another instance of the confused boundary between bibliographic and authority control records can be seen in the handling of a name variation for a title of a bibliographic item . . . the practice persists of including title variations for a bibliographic item in bibliographic records. Logically, they are name variations for the work and belong in an authority record."

Granted that object-oriented modeling is likely to be a powerful approach to library computer implementations, why should it not simply be adopted for use with existing cataloging standards and rules? What relevance does it have for the wider cataloging world?

One answer to this question can be found in the work of Gregory H. Leazer, who has analyzed the redundancy inherent in MARC records. The entity-relationship model underlies Leazer's thinking. He concludes:

The conceptual schema of the current USMARC formats is inadequate, and the structure is confused. . . . The conceptual schema also should include justifications for each entity and relationship included in the database. For each data entity and relationship, specifications for their inclusion should be articulated and their organization specified. (Leazer 1992, 204)

I have already intimated that the current bibliographic standards intermix bibliographic details from

different levels. If our approach to cataloging changes, we can more easily construct records that mirror natural generalizations about the bibliographic world. For example, an obvious concept might be: "all of the editions of Hamlet." At present queries such as these tend to be approached via title searches on "Hamlet," possibly with "Shakespeare" as a name heading. Very few systems, if any, carry the concept of "Hamlet" as a work with a particular identity, linked to "Shakespeare" as author independently of any bibliographic records.

Bob Strauss, of the University of North Carolina at Asheville, offered a telling example of this on the Auto-cat discussion list in December 1993. He cited the editions of Ed Krol's *Whole Internet Catalog* on OCLC:

There are 9 entries on OCLC, as of today, 12/2/93. Number one, is DLC's. I am listing them, without "credit" but with the holdings attached, and the "differentiating" feature. The search I used was: krol,whol: (Strauss 1993).

Number	Holdings	Feature
1 (DLC DLC)	323	1992
2	10	minor corr, 1992
3	883	[Corr ed]
4	968	1st ed (same as DLC?)
5	51	July 1993, minor corr.
6	136	"May 1993"
7	116	[Corr ed] (1993)
8	19	[Corr ed]
9	132	"Feb. 1993; minor corr."

Total: 2,638 copies

Strauss continues:

We all "know" that this is really one work. Since it is electronic, it is easy to revise. So each succeeding printing has a few revisions. The problem is that, as this sort of thing becomes more common, we will end up with many records for thousands of works. How many Krol volumes will there be in 5 years?

The implementation of the Z39.50 protocol, with the promise of searching several databases simultaneously, makes the introduction of the "one work" concept even more desirable.

Another concept that many systems handle poorly is, surprisingly enough, the concept of the physical copy of the book. Copy numbers and call numbers are handled, but less well handled are, for example, manu-

script annotations in a book, notes about bookplates, or the record of which number of a limited edition numbered set a library owns. It is not easy to bring together all of the information relating to single copies of books because the MARC record handles them in different parts of the record or, in some cases, does not handle them at all. In a record-sharing environment there are even advantages to disregarding this sort of information.

Modeling the Cataloging World

First Principles

How then should we go about modeling the cataloging world? First it is perhaps necessary to repeat some basic facts about object-oriented modeling. Object-oriented modeling is not a database system and has nothing directly to do with computer programming or programming languages. It is simply a conceptual tool to assist in thinking about objects in a very broad sense, as well as their relations and the events in which they participate. Anything can be an object. Objects have attributes and have a known set of operations in which they can participate. They can be grouped together in classes that share sets of attributes and operations, and they can be aggregated together. Objects, and classes of objects, can be associated in various ways. Very importantly, objects can pass on their characteristics to other objects with which they have some sort of relation, and the general attributes of classes are shared by the objects belonging to that class.

In most applications of object orientation, the object-oriented model is used to simulate situations in which events occur. For example, in the personnel management of a company in which there are persons, jobs, salaries, workplaces, and so on, the purpose of the model is to mimic real-world events which take place to change the relations among these elements: the appointment of a new worker, a promotion, a change of workplace, or whatever. While there are obvious parallels to this in the library world—for example, in tracking the transactions that take place between borrowers, books, and loan periods—I want to concentrate on a somewhat different application. In the model I will describe, we will not be concerned so much with transactions as with a comparatively static model of relations between authors, titles, publishers, and the like, in which the events are the generation of one object from another by translation, revision, or other means. Once an event has taken place, it cannot be reversed: a book, for example, cannot be "unpublished."

The following is a very simple example of object orientation: We can define there to be a class of objects called "Persons." Persons have particular attributes, such as a name, a sex, a clothing size. Mary Smith is an object who is an instance of the class of Persons. She inherits the attributes we have listed and assigns particular values to each of them. If the class Persons can get married or buy clothing, so can Mary Smith. Mary Smith may enter into relations with other objects. If there is a class of objects called Cars, then Mary Smith can be in an "ownership" relation with an instance of that class, or (if she owns two cars) with more than one instance.

The key concepts are the natural ones: that things ("objects") exist, that we can group things into classes and describe in general terms the relevant facts that characterize things (their "attributes"), and the sort of things that they may do, or that may happen to them (their "operations"). Different things may be involved with or linked to each other in different ways (their "relationships"). At this level of description, our main concerns are to deal with these concepts in a naturalistic manner, without getting bogged down in details of implementation (such as assignation of unique reference numbers) and to look at the concepts pragmatically—i.e., with an eye to the practical purposes for which the model is intended, rather than to the construction of a grand philosophical representation of life, the universe, and everything else.

The Text Model

Now it is time to take a look at a similar description of something related to bibliographic records. This overview will be simplistic, of course, but can serve as a starting point from which more complex situations could be approached.

There is a class of objects which we shall call "Texts." Texts are strings of sentences. They have titles and they are in a language. They can be subjected to various operations, including translation, revision, and aggregation with other texts. Each instance of the class will have values for these attributes; and there are derived classes of derived instances which may have other attributes, as well.

This can be represented schematically (figure 1).

"A Christmas Carol" is an object that is an instance of the class (figure 2).

From this object another object may be derived (figure 3).

The two objects are associated with each other by the operation of translation. Once this operation has taken place, they remain linked. The attributes of original title and language can be "copied" into the derived

```

Class: Text
Attributes: title
           language
Operations: translation
           revision
           aggregation
           ...

```

Figure 1a

```

Class: Derived text
Attributes: title
           language
           derived title
           derived language
Operations: translation
           revision
           aggregation
           ...

```

Figure 1b

```

Instance
of class: Text
Attributes: title: Christmas carol
           language: English
Operations: translation: no
           revision: no
           aggregation: no
           ...

```

Figure 2

```

Instance
of class: Derived text
Attributes: title: Christmas carol
           language: English
           derived title: Weihnachtslied
           derived language: German
Operations: translation: yes
           revision: no
           aggregation: no
           ...

```

Figure 3

text, or they may be left implied by their existence at the other end of the link.

There is another class of objects, "Agents," which may be associated with Texts. Each Text may be associated with several Agents, and each Agent may be associated with several Texts. Agents may be authors, editors, revisers, translators, and so on. The precise nature of the association can be modeled as a link attribute; that is, the attribute belongs to the link, not to the Agent or the Text. Agents have attributes such as names or dates of birth, if they are persons; if they are corporate bodies, then there are agents and subagents, such as the departments of a university. Agents associated with primitive Text objects are usually inherited by derived Text objects.

The Publication Model

I have described, very simplistically, a model for texts and the persons associated with them. All of this is done without reference to any published work. Published works are manifestations of the texts. A particular text object can be embedded in a publication. "Publication" is a class which has its own attributes and operations. The overall class of publications can be divided into several media (books, films, etc.) (figure 4).

The class of Agents also has associations with this class of Publications—as publishers, printers, distributors, and so on.

The operations in which publications can participate are of two main types: reformatting, which can transform any or all of a publication's attributes and cause it to change from one subclass into another; and republication, which may involve reformatting, but also involves a change of Agent. Reformatting may be as simple as the issue of a paperback from a hardback; or it may involve a major transformation, such as the production of a CD-ROM from the hard-copy text. Republication may be, for example, the production of a facsimile or of a microfiche edition by a different publisher.

I have omitted date from the characteristics of both texts and publications. Dates may be associated with both the text and the publication. In the nature of things, the date of a text will be more elusive than the date of publication. In some cases it becomes important to know or to specify the date of composition of a text, as well as its date of publication. In most circumstances, however, people refer to the date of publication both as such and as a surrogate for the date of composition, so on pragmatic grounds it is better to view date as an attribute of the Publication model. On the other hand, some works are published on demand and, for these, the idea of a date of publication may not be suitable. As with most

```

Class: Publication
Subclass: Book
Attributes: extent: length
           dimension: size
           packaging: binding
           identification number
Operations: reformatting
           republication

```

Figure 4

```

Class: Copy
Attributes: call number
           location
           binding quality
           ...
Operations: loan
           reservation
           sending for binding
           ...

```

Figure 5

aspects of object-oriented modeling, decisions should be made on pragmatic grounds.

Other obvious areas for debate are edition and series. Editions sometimes designate new texts and sometimes designate only publishers' changes. Similarly, series have titles but may perhaps be more closely associated with publishers than with authors. In the modular structure that I am describing, the pragmatic approach would be to draw the boundaries between the models in such a way as to minimize the connectivity between them.

The Copy Model

The Publication model describes the generality of a particular publication. For most forms of publication, each publication is manifested in a certain number of physical copies. Every copy has its own characteristics. It embodies all of the characteristics inherited from the publication model, which in turn embodies those inherited from the text model. In describing a copy in a library (I shall disregard here the other situations in which copies of books exist), there are several

T 090 \$a PR3724 \$b .G7 1726b
T 100 10 \$a Swift, Jonathan, \$d 1667-1745
T 240.10 \$a Gulliver's travels
P/T 245 10 \$a Travels into several remote nations of the world. : \$b In four parts. / \$c By Lemuel Gulliver, first a surgeon, and then a captain of several ships ; with annotations by Daniel Defoe
P 260 \$a London : \$b Printed for Benj. Motte, at the Middle Temple-Gate in Fleet-street., \$c MDCCXXVI
P 300 \$a 2 v. , VI pl. : \$b 5 maps, 1 diagr. ; \$c 19 cm
T 500 \$a By Jonathan Swift
P 500 \$a Each part has separate paging, and pts. I-II and IV, special title-pages; pt. III has only general t.-p
P 500 \$a Apparently the second edition. ("AA" in H. Teerink's A bibliography of the writings of Jonathan Swift. 1963. p. 192-198)
C 500 \$a Bodleian Library's copy has a bookplate of the Benedictine Monastery in Graftschafft; inscribed by Charles Dickens \$5 BOD
T 650 0 \$a Voyages, Imaginary
T 700 10 \$a Defoe, Daniel
C 700 10 \$a Dickens, Charles \$4 ins \$5 BOD
P 700 10 \$a Motte, Benjamin \$4 pbl
C 710 20 \$a Graftschafft (Monastery) \$4 fmo \$5 BOD

Figure 6

specific attributes and operations for which provision should be made (figure 5).

The copy may have associations with Agents in several ways: it has an owning Agent, possibly previous owning Agents, and in a library may be associated with a borrowing or reserving Agent. The physical copy may be aggregated with the copies of other publications, as in the case of bound sets of pamphlets.

Subjects and Synonyms

So far I have said nothing about subject description. However, subject and classification can clearly be associated with texts, and could be added into the model on that basis.

Besides the thesaural set of subject terms, subject headings can also relate to the names of persons, works, or even editions of works (such as a work about the first quarto of Hamlet). This means that in the object modeling environment, various objects can themselves function as subjects.

Subject and thesaural systems usually have well-developed syndetic structures of their own to cope with synonyms and related terms. Similarly, name authority files have such structures. There is no reason why a similar set of synonyms should not be created for titles to handle cover titles, spine titles, and the like. Note, however, that in the above analysis of texts, a part title designates a part text, and the difference in text should be modeled in the text model.

Mapping Object-Oriented Cataloging to MARC

If we apply considerations of the sort outlined above to an AACR2/MARC record, it is clear that the three models we have defined are jumbled up within it. In order to explore all of the models, in the record below we shall incorporate the elements of US MARC for rare book cataloging, including library codes and name relators. Figure 6 is based upon a real record but has had extra fields added to demonstrate the three models. The letters at the left indicate whether the field in question belongs to the text model (T), the publication model (P), or the copy model (C).

If we rearrange the record into its constituent models, we get the following (figure 7).

In the United Kingdom, the UBIS Working Group on Bibliographic Standards has been similarly exercised by the problems of the technical limitations of the UK MARC format, bibliographic relations, and the like.

Colin Lunt gives two examples of the complexities involved: one of them is Douglas Adams's *Hitchhiker's Guide to the Galaxy* (Lunt 1993). Here (figures 8-14) is a set of US MARC records (some in outline only) describing different parts of and/or different editions of the *Hitchhiker's Guide*.

In an object-oriented universe, these might be handled as follows (figure 15).

(Text continues on page 146)

T 090 \$a PR3724 \$b .G7 1726b
T 100 10 \$a Swift, Jonathan, \$d 1667-1745
T 240 10 \$a Gulliver's travels
T 245 10 \$a Travels into several remote nations of the world
T 500 \$a By Jonathan Swift
T 650 0 \$a Voyages, Imaginary
T 700 10 \$a Defoe, Daniel

Figure 7a

P 245 10 \$a Travels into several remote nations of the world. : \$b In four parts. / \$c
By Lemuel Gulliver, first a surgeon, and then a captain of several ships ; with
annotations by Daniel Defoe
P 260 \$a London : \$b Printed for Benj. Motte, at the Middle Temple-Gate in
Fleet-street., \$c MDCCXXVI
P 300 \$a 2 v. , VI pl. : \$b 5 maps, 1 diagr. ; \$c 19 cm
P 500 \$a Each part has separate paging, and pts. I-II and IV, special title-pages;
pt. III has only general t.-p
P 500 \$a Apparently the second edition. ("AA" in H. Teerink's A bibliography of
the writings of Jonathan Swift. 1963. p. 192-198)
P 700 10 \$a Motte, Benjamin \$4 pbl

Figure 7b

C 500 \$a Bodleian Library's copy has a bookplate of the Benedictine Monastery in
Grafschaft; inscribed by Charles Dickens \$5 BOD
C 700 10 \$a Dickens, Charles \$4 ins \$5 BOD
C 710 20 \$a Grafschaft (Monastery) \$4 fmo \$5 BOD

Figure 7c

...
020 \$a 0330258648
040 \$a GZM \$c GZM \$d m/c \$d BGU
049 \$a EQQA
090 \$a PR6051.D3352 \$b H57 1979b
100 10 \$a Adams, \$j Douglas, \$d 1952-
245 14 \$a The hitch hiker's guide to the galaxy / \$c Douglas Adams
260 \$a London : \$b Pan Books, \$c 1979
300 \$a 158 p. ; \$c 18 cm
500 \$a "A Pan original"--Cover
650 0 \$a Radio serials
650 0 \$a Voyages, Imaginary

Figure 8

...
245.14 \$a The hitch-hiker's guide to the galaxy. \$n Part 2, \$p The restaurant at the end of
the universe / \$c by Douglas Adams
260 \$a London \$b Original Records \$c c1980
300.00 \$a 1 sound cassette
...

Figure 9

...
020 \$a 0671739670 (pbk.) : \$c \$4.95
040 \$a APL \$c APL \$d HQB \$d OCL
100 10 \$a Adams, \$j Douglas, \$d 1952-
245 10 \$a Life, the universe and everything
/ \$c Douglas Adams
260 \$a New York : \$b Pocket Books,
\$c 1983, c1982
300 \$a 227 p. ; \$c 18 cm

Figure 10

010 \$a LC 84-019350 //r92
020 \$a 0517554399 : \$c \$12.95
040 \$a DLC \$c DLC \$d OCL \$d IEP \$d SVP
\$d TXA
050 00 \$a PR6051.D3352 \$b S6 1985
082 00 \$a 823/.914 \$2 19
100 10 \$a Adams, \$j Douglas, \$d 1952-
245 10 \$a So long, and thanks for all the
fish / \$c Douglas Adams
250 \$a 1st ed
260 \$a New York : \$b Harmony Books,
\$c c1985
300 \$a 204 p. ; \$c 20 cm

Figure 11

010 \$a LC gb- 85039886
015 \$a GB85-39886
020 \$a 0330287001 (pbk) : \$c f1.95
040 \$a UKM \$c UKM \$d m/c
082 0 \$a 823/.914 \$2 19
100 10 \$a Adams, \$j Douglas, \$d 1952-
245 10 \$a So long, and thanks for all the
fish : \$b the hitchhikers' guide
to the galaxy 4 / \$c Douglas Adams
260 \$a London : \$b Pan, \$c 1984 \$g (1985
[printing])
300 \$a 191 p. ; \$c 18 cm

Figure 12

010 \$a LC 86-019480 //r94
020 \$a 0517564254
040 \$a DLC \$c DLC \$d JED
050 00 \$a PR6051.D3352 \$b H54 1986
082 00 \$a 823/.914 \$2 19
100 10 \$a Adams, Douglas, \$d 1952-
245 14 \$a The hitchhiker's quartet / \$c Douglas Adams
250 \$a 1st ed
260 \$a New York : \$b Harmony Books, \$c c1986
300 \$a xi, 624 p. : \$b ill. ; \$c 24 cm
505 0 \$a Introduction : a guide to the guide - The hitchhiker's guide to the galaxy - The
restaurant at the end of the universe - Life, the universe and everything - So long,
and thanks for all the fish - Young Zaphod plays it safe

Figure 13

010 \$a LC gb- 92359791
015 \$a GB92-Z9791
020 \$a 0434009261 : \$c f13.99
040 \$a UKM \$c UKM \$d CUY
082 04 \$a 823.914 \$a F \$2 20
090 \$a PR6051.D3352 \$b M7 1992b
100 10 \$a Adams, Douglas, \$d 1952-
245 10 \$a Mostly Harmless / \$c Douglas Adams
260 \$a London : \$b Heinemann, \$c 1992
300 \$a 219 p ; \$c 21 cm

Figure 14

[Text object 1]
[Class: Primitive text object]
[Title] Hitch-hiker's guide to the galaxy
(1979)
[Association: synonym] Hitch-hiker's guide
to the galaxy
[Association: author: Agent object 1]
[Association: number n in list]: 1
[Association: nth component of: Text object 7]

Figure 15a

[Text object 2]
[Class: Primitive text object]
[Title] Restaurant at the end of the
universe
[Association: author: Agent object 1]
[Association: number n in list]: 2
[Association: nth component of: Text object 7]

Figure 15b

[Text object 3]
[Class: Primitive text object]
[Title] Life, the universe and everything
[Association: author: Agent object 1]
[Association: number n in list]: 3
[Association: nth component of: Text object 7]

Figure 15c

[Text object 4]
[Class: Primitive text object]
[Title] So long, and thanks for all the fish
[Association: author: Agent object 1]
[Association: number n in list]: 4
[Association: nth component of: Text object 7]

Figure 15d

[Text object 5]
[Class: Primitive text object]
[Title] Young Zaphod plays it safe
[Association: author: Agent object 1]
[Association: may be 3rd component of: Text object 8]

Figure 15e

[Text object 6]
[Class: Primitive text object]
[Title] Introduction : a guide to the guide
[Association: author: Agent object 1]
[Association: may be 1st component of: Text object 8]

Figure 15f

[Text object 7]
[Class: Derived text object]
[Title] Hitch-hiker's guide to the galaxy
(1992)
[Association: synonym] Hitch-hiker's guide to the galaxy

Figure 15g

[Text object 8]
[Class: Derived text object]
[Title] Hitch-hiker's quartet
[Language English]
[Association: has as 1st component: Text object 6]
[Association: has as 2nd component: Text object 7]
[Association: has as 3rd component: Text object 5]

Figure 15h

[Text object 9]
[Class: Primitive text object]
[Title] Mostly harmless
[Association: author: Agent object 1]

Figure 15i

[Agent object 1]
[Name] Adams, Douglas
[Agent type Corporation]
[Attribute]: b.1952

Figure 15j

[Agent object 2]
[Name] Harmony Books
[Agent type Corporation]
[Attribute: Place]: New York

Figure 15k

[Agent object 3]
[Name] Pocket Books
[Agent type Corporation]
[Attribute: Place]: New York

Figure 15l

[Agent object 4]
[Name] Original Records
[Agent type Corporation]
[Attribute: Place]: London

Figure 15m

[Agent object 5]
[Name] Pan
[Agent type Corporation]
[Attribute: Place]: London

Figure 15n

[Agent object 6]
[Name] Heinemann
[Agent type Corporation]
[Attribute: Place]: London

Figure 15o

[Publication object 1]
[Class: Primitive publication object]
[Subclass: Book]
[Content: Text object 1]
[Extent] 158 p.
[Dimensions] 18 cm
[Packaging] pbk
[Date] 1979
[ID number]0330258648
[Association: publisher: Agent object 5]

Figure 15p

[Publication object 2]
[Class: Derived publication object]
[Subclass: Sound recording]
[Content: Text object 2]
[Extent] 1 sound cassette
[Dimensions]
[Packaging]
[Date] 1980
[ID number]
[Association: publisher: Agent object 4]

Figure 15q

```
[Publication object 3]
[Class: Primitive publication object]
[Subclass: Book]
[Content: Text object 3]
[Extent] 227 p.
[Dimensions] 18 cm
[Packaging] pbk
[Date] 1983
[ID number]0617739670
[Association: publisher: Agent object 3]
```

Figure 15r

```
[Publication object 4]
[Class: Primitive publication object]
[Subclass: Book]
[Content: Text object 4]
[Extent] 204 p.
[Dimensions] 20 cm
[Packaging] pbk
[Date] 1985
[ID number]0517554399
[Association: publisher: Agent object 2]
```

Figure 15s

```
[Publication object 5]
[Class: Primitive publication object]
[Subclass: Book]
[Content: Text object 4]
[Extent] 191 p.
[Dimensions] 18 cm
[Packaging] pbk
[Date] 1984
[ID number]0330287001
[Association: publisher: Agent object 5]
```

Figure 15t

(Text continues from page 142)

It will be clear that concepts such as the construction of a uniform title or an author's name are being used just as they are in AACR2; and that many MARC codings have a direct translation into an object-oriented cataloging model. Equally, cataloging judgments have to be made about the relationships among texts and among publications.

The title proper in the 245 field of a MARC record is at the same time a variant of the title and, with the rest of the title field, an attribute of the publication. The full title field, if it is to be used, belongs in the publication object rather than in the text object.

I have not elaborated further here on the concept of the copy model. I hope it is clear that the core would be the embedding of the publication object in a record

```
[Publication object 6]
[Class: Primitive publication object]
[Subclass: Book]
[Content: Text object 8]
[Extent] xi, 624 p.
[Dimensions] 24 cm
[Packaging] hbk
[Date] 1986
[ID number]0517564254
[Association: publisher: Agent object 2]
```

Figure 15u

```
[Publication object 7]
[Class: Primitive publication object]
[Subclass: Book]
[Content: Text object 9]
[Extent] 219 p.
[Dimensions] 21 cm
[Packaging] hbk
[Date] 1992
[ID number]0434009261
[Association: publisher: Agent object 6]
```

Figure 15v

that would also have the attributes and the ability to participate in events as described in the generalized model given above. Another conceptualization would be to say that the objects of the publication model become the classes of the copy model: each copy is an instance of its class, the published work.

The Uniform Title Authority File

If we try to imagine how to implement object-oriented cataloging, it is clear that, while much could be carried over with little change, various developments would have to take place moving us away from current MARC and AACR2 and towards the erection of such models as I have indicated above. In current terms, a first stage would be the creation of a uniform-title authority file and the automatic assignation of uniform titles to all works, not just to those which are perceived to "need" them. This would be the maturation of a process that has been growing for the better part of 150 years from Panizzi's construction of uniform headings for certain works. In theory a uniform-title authority file should be separate from a name authority file, but in practice it may be possible to develop one out of the other, with uniform titles as subfields in an author record, or with repeated author records identifying each uniform title,

or by means of the introduction of a more complex relational structure within the authority file to allow both names and titles to be present only once, with pointers between them.

Most of the ideas above have been expressed before. The core of the proposed re-analysis—and the single most significant shift—is the proposal that an authority file of titles be created as a means of recognizing the “work” as the core entity with which we are concerned, that this file be associated with the name authority file, and that the titles authority file should then point to bibliographical items. This idea, precisely, was expressed as long ago as 1977 by Michael Gorman in a conference reviewing the first decade of automated cataloging (Gorman 1979). What is new here, perhaps, is that in the period since Gorman spoke computing tools have caught up with his ideas. Object-oriented modeling is the general conceptual tool by which his suggestions can be implemented, and computer programs of general application are now being constructed to embody such modeling.

Getting There from Here

If the model outlined above is an accurate description of the target structures towards which we should be aiming, there are three aspects to the problem of achieving that aim. These are:

- the construction of linked authority files reflecting the classes of objects;
- the refinement of MARC records so that they may both remain hospitable to the present formats and be carriers of the new layers of information; and
- solving the problem of intertwining national (or transnational) authority files—embodying much of the knowledge-base of the bibliographical world—with local systems containing subsets of those files but also carrying other or related objects not present in them.

Linked Authority Files

The USMARC authority record structure is probably the best developed linked authority file, and it is the one I shall consider here. In the Library of Congress Name Authority File (LC NAF) each record has a number. Unauthorized forms are listed as tagged fields in the record whose “main” field is the authorized form. For simple names, each record can be seen as a primitive object. For corporate names with subfields, some deconstruction has to take place to produce a parent record

with offspring in the form of departments, etc. Parent and offspring need a linking field:

“this-department : is-a-part-of : this-object”

In the object model, “is-a-part-of” is an attribute of the association rather than of either of its objects. However, for the practical purposes of moving from existing structures, it can be implemented within either object participating in the association or within both. In other words, in a MARC authority record based on object orientation, “is-a-part-of” can be joined either to “this-department” or to “this-object,” although it actually belongs to neither.

```
010      n 79095521
040      $a DLC $c DLC $d DLC
110 20   $a University of Oxford
410 10   Oxford. $b University
410 20   Prifysgol Rhydychen
410 20   Oxford University
410 20   Academia Oxoniensis
410 20   J_ami_at Uksf_urd
670      $a The Spirit lamp, 1892-
670      $a Car use, c1983 (subdiv.) $b CIP
         t.p. (Oxford University)
670      $a InU/Wing STC files $b (usage:
         ... Academia Oxoniensi)
```

Figure 16

```
010      n 80020306
040      $a DLC $c DLC $d DLC $d DLC-R
110 20   $a Bodleian Library
410 10   $a Oxford. $b University.
         $b Bodleian Library
410 20   $a University of Oxford.
         $b Bodleian Library
410 20   $a Bibliotheca Bodlejiana
410 20   $a Sifriyat Bodli
410 20   $a Bodleian Library
410 20   $a Bodleyanah
410 20   $a Sifriyat Bodleyanah
670      $a Its English music, 1955
```

Figure 17

```
010      n 78099222
040      $a DLC $c DLC $d MnHi
110 20   $a University of Oxford. $b Faculty
         of Music
410 10   $a Oxford. $b University.
         $b Faculty of Music
670      $a Its The Bate collection ...
         1976: $b t.p. (Faculty of Music,
         University of Oxford)
```

Figure 18

Take as an example the following LC NAF records (figures 16–18) for the University of Oxford and two of its institutions. If we introduce a relator field (say, 666) with relator codes (say “pt” for “is-a-part-of”), the records for the institution can be revised as in figure 19a (or more radically, as in figure 19b, in which a field 210 is used to indicate a body that needs its parent to create the full record) and 20.

- 210 to indicate this form cannot be used alone and requires its 666 parent
- 110 to indicate this form can be used alone and does not require its 666 parent

In a migration to such an arrangement, the new field(s) could be added to the standard without any

```
010    n 78099222
040    $a DLC $c DLC $d MnHi
110 20 $a University of Oxford. $b Faculty
      of Music
410 10 $a Oxford. $b University.
      $b Faculty of Music
666 00 $a pt $n n 79095521
670    $a Its The Bate collection ...
      1976: $b t.p. (Faculty of Music,
      University of Oxford)
```

Figure 19a

```
010    n 78099222
040    $a DLC $c DLC $d MnHi
210 20 $b Faculty of Music
410 10 $a Oxford. $b University.
      $b Faculty of Music
666 00 $a pt $n n 79095521
670    $a Its The Bate collection ...
      1976: $b t.p. (Faculty of Music,
      University of Oxford)
```

Figure 19b

```
010    n 80020306
040    $a DLC $c DLC $d DLC $d DLC-R
110 20 $a Bodleian Library
410 10 $a Oxford. $b University.
      $b Bodleian Library
410 20 $a University of Oxford.
      $b Bodleian Library
410 20 $a Bibliotheca Bodlejiana
410 20 $a Sifriyat Bodli
410 20 $a Bodleian Library
410 20 $a Bodleyanah
410 20 $a Sifriyat Bodleyanah
666.00 $a pt $n n 79095521
670    $a Its English music, 1955
```

Figure 20

restructuring. Some fields and subfields would become essentially redundant, but would be retained for backward compatibility.

Uniform titles are already present for works without authors in the LC NAF, together with name-title added entries. Name-title entries could be handled in a manner analogous to that suggested for subordinate bodies for corporate names, although in this case there would be a greater premium on having titles entirely independent of their authors, as this would match more closely the object model.

The extra requirement is for linkage between authors and their titles. If the current usage for name-title entries is retained, then a linkage is automatically

```
010    $a n 79061368
040    $a DLC $c DLC $d MdU $d DLC-R $d DLC
053    $a PQ2605.A3734
100 10 $a Camus, Albert, $d 1913-1960
400 10 $a Kam_i_u, Al'ber, $d 1913-1960
...
670    $a His The stranger, 1946.
670    $a Y_uan, S.C. K_a-miu ti huang miu
      che hs_ueh, 1989.
```

Figure 21

```
010    $a n 50027540 $z n 92088761
040    $a DLC $c DLC $d NIC $d DLC
100 10 $a Gilbert, Stuart
670    $a Joyce, J. $b Ulysse ... 1929.
670    $a His Recollections of James
      Joyce, 1993: $b CIP t.p. (Stuart
      Gilbert) pub. info. (b. 1883;
      d. 1969)
```

Figure 22

```
[Title Authority record\]
010    $a t 12345678
040    $a BOD $b BOD
041.10 $a fre
130.00 $a Peste
667 00 $a n 79061368 $4 aut
```

Figure 23

```
[Title Authority record\]
010    $a t 23456789
040    $a BOD $b BOD
041.10 $a fre $h eng
240.00 $a Plague
667.00 $a t 12345678
667.00 $a n 500275400 $4 trl
```

Figure 24

present, albeit with massive redundancy. If titles are established essentially independent of names, then a linking field is required. Using current LC NAF records for names and using a "model" record for the posited Title Authority File, the following authority records (figures 21-24) could be used in a bibliographical record.

Refinement of MARC Records

Bibliographical records could be augmented by codes at the end of each relevant field, pointing to the appropriate authority record (and, if necessary, the authority file from which it comes). The first step in such a process may produce a record such as this (figure 25).

This record contains apparently independent links from name to name authority file and from title to title authority file. In a fully developed object-oriented system, the names would be derived from the title identifier. The publisher would also have a link to the Names Authority file. In current MARC practice, publisher access would be by an added 7XX field, which could then be linked to the NAF. In a more fully fledged implementation, however, the link should be contained directly in the 260 field.

Eventually it should prove possible in any given system to dispense with the text of the linked field and to retain only the link identifier. The key to this development is the ability to relate local databases to external systems.

As an example let us return to the example given earlier which, sorted to its component models, produced figure 26.

In a fully developed object-oriented cataloging system, the component records might appear as in figure 27.

I should stress that these examples are given only to demonstrate that there are ways of adapting MARC so that it reflects more closely an object-oriented analysis. It is an implementation within a given framework,

```

008   $b r $c 1991 $d 1948 $e nyu $q 0 $r
      0 $s 0 $v 1 $x eng
010   $a LC 90-050477 //r91
020   $a 0679720219 (pbk.)
040   $a DLC $c DLC
041 1  $a eng $h fre
050 00 $a PQ2605.A3734 $b P413 1991
082 00 $a 843/.914 $2 20
100 10 $a Camus, $j Albert, $d 1913-1960 $2
      LC n 79061368
240 10 $a Peste. $1 English $2 LC t 23456789
245 14 $a The plague / $c Albert Camus ;
      translated from the French by
      Stuart Gilbert
250   $a 1st Vintage international ed
260   $a New York : $b Vintage Books,
      $c 1991
300   $a 308 p. ; $c 21 cm
500   $a Translation of: La peste
700 10 $a Gilbert, Stuart $2 LC n 50027540
  
```

[Another translation would be a different title "object" and so the 240 would have a different reference number; the 041 field is of course also copied from/copiable to the authority record.]

Figure 25

```

T 090   $a PR3724 $b .G7 1726b
T 100 10 $a Swift, Jonathan, $d 1667-1745
T 240.10 $a Gulliver's travels
T 245 10 $a Travels into several remote
      nations of the world
T 500   $a By Jonathan Swift
T 650 0  $a Voyages, Imaginary
T 700 10 $a Defoe, Daniel
  
```

Figure 26a

```

P 001   $a OCLC 11392229
P 245 10 $a Travels into several remote nations of the world. : $b In four parts. /
      $c By Lemuel Gulliver, first a surgeon, and then a captain of several ships ; with
      annotations by Daniel Defoe
P 260   $a London : $b Printed for Benj. Motte, at the Middle Temple-Gate in
      Fleet-street., $c MDCCLXXVI
P 300   $a 2 v. , VI pl. : $b 5 maps, 1 diagr. ; $c 19 cm
P 500   $a Each part has separate paging, and pts. I-II and IV, special title-pages; pt.
      III has only general t.-p
P 500   $a Apparently the second edition. ("AA" in H. Teerink's A bibliography of the
      writings of Jonathan Swift. 1963. p. 192-198)
P 700 10 $a Motte, Benjamin $4 pbl
  
```

Figure 26b

C 500 \$a Bodleian Library's copy has a bookplate of the Benedictine Monastery in Graftschafft; inscribed by Charles Dickens \$5 BOD

C 700 10\$a Dickens, Charles \$4 ins \$5 BOD

C 710 20 \$a Graftschafft (Monastery) \$4 fmo \$5 BOD

Figure 26c

Title Authority File [TAF]:

010 \$a t 23456789

041.10 \$a eng

090 \$a PR3724 \$b .G7 1726b

130.00 \$a Gulliver's travels

430.00 \$a Travels into several remote nations of the world

667 00 \$a n 78096912 \$2 LCNAF \$4 aut

667.00 \$a n 79053974 \$2 LCNAF \$4 ann

667 00 \$a sh 85144441 \$2 LCSH

[The generalized elements of the classification should be derived from the appropriate authority file]

Figure 27a

NAF

010 \$a n 78096912

040 \$a DLC \$c DLC \$d DLC \$d OCoLC \$d MdU \$d DLC

053 \$a PR3720 \$b PR3728

100 10 \$a Swift, Jonathan, \$d 1667-1745 [AACR2]

400 10 \$a Gulliver, Lemuel, \$d 1667-1745

400 10 \$a Swift, Dzhonatan, \$d 1667-1745

400 10 \$a Du Baudrier, \$c sieur, \$d 1667-1745

400 10 \$a Wagstaff, Simon, \$d 1667-1745

...

670 \$a His Some free thoughts upon the present state of affairs, 1741: \$b t.p. (Author of Gulliver's travels)

Figure 27b

NAF

010 \$a n 79053974

040 \$a DLC \$c DLC \$d DLC \$d MdU \$d DLC

053 \$a PR3400 \$b PR3408

100 10 \$a Defoe, Daniel, \$d 1661?-1731 [AACR2]

400 10 \$a Defo, Daniel, \$d 1661?-1731

400 00 \$a Gabriel John, \$d 1661?-1731

...

400 10 \$a Moreton, Andrew, \$d 1661?-1731

400 00 \$a Gentleman, \$d 1661?-1731

400 00 \$a L. M., \$d 1661?-1731

...

670 \$a His A tour thro' the whole island of Great Britain, 1724-1727: \$b t.p. (a gentleman)

...

670 \$a Halkett & Laing \$b (L.M., Daniel Defoe)

Figure 27c

NAF

010 \$a n 78087607

040 \$a DLC \$c DLC \$d DLC \$d MdU \$d DLC \$d IU

053 \$a PR4550 \$b PR4598

100 10 \$a Dickens, Charles, \$d 1812-1870 [AACR2]

400 10 \$a Boz, \$d 1812-1870

400 10 \$a Dickens, Karol, \$d 1812-1870

...

670 \$a Libman, Z._I.A. Charlz Dikkens, 1982: \$b t.p. (Charlz Dikkens)

670 \$a Page, N. Bleak House, c1990: \$b CIP galley (Charles John Huffam Dickens)

670 \$a His _Syk_stuolis Skrud_zas, 1922: \$b t.p. (_Carlzas Dikensas)

Figure 27d

NAF

010 n 80020306

040 \$a DLC \$c DLC \$d DLC \$d DLC-R

110 20 \$a Bodleian Library

410 10 \$a Oxford. \$b University. \$b Bodleian Library

...

Figure 27e

NAF

[constructed\]

010 \$a n 87654321

040 \$a DLC \$c DLC

110 20 \$a Graftschafft (Monastery)

Figure 27f

NAF

```
[constructed\  
010 $a n 98765432  
040 $a DLC $c DLC  
100 10 $a Motte, Benjamin  
...  
265 00 $a London, Fleet Street, Middle  
Temple Gate
```

Figure 27g

SAF

```
010 $a sh 85144441  
040 $a DLC $c DLC  
053 $a G560  
150 0 $a Voyages, Imaginary  
450 0 $a Imaginary travels  
450 0 $a Imaginary voyages  
450 0 $a Travels, Imaginary  
550 0 $a Exoticism in literature  
550 0 $a Robinsonades  
550 0 $a Voyages to the otherworld
```

Figure 27h

not my conception of how object-oriented cataloging could best be structured. The links, for example, are embedded in one end of each relationship: so that there is a direct link from the record for the published text of *Gulliver's Travels* to Swift as author, but no direct link from the record for Swift to *Gulliver's Travels*. A sounder implementation would also have pointers in the record for Swift to *Gulliver's Travels* and his other works, or a join table for the lists of names and titles.

Intertwining Local and External Systems

An object-oriented cataloging universe can only reach full development if everyone agrees on the identifiers to be used in the creation of objects. Records consisting mainly of strings of commonly recognized identifiers could then be exchanged. Until then some form of duplication, as suggested above, would be needed. The general consensus is already present for names and subject headings in the form of the LC authority files; and, for bibliographic manifestation records, the LC or OCLC number could form the core of such an agreement.

There are some interrelated developments that make the Ultimate Cataloging Universe more than just a pipe dream. The first of these is the development of the Internet and of high-speed links between machines,

```
001 $a OCLC 11392229  
240 00 $a t 23456789  
245 10 $a Travels into several remote  
nations of the world. : $b In four parts.  
/ $c By Lemuel Gulliver, first a surgeon,  
and then a captain of several ships ;  
with annotations by Daniel Defoe  
260 $a n 98765432 $4 pbl $c MDCCXXVI  
300 $a 2 v. , VI pl. : $b 5 maps, 1  
diagr. ; $c 19 cm  
500 $a Each part has separate paging,  
and pts. I-II and IV, special  
title-pages; pt. III has only general t.-p  
500 $a Apparently the second edition.  
("AA" in H. Teerink's A bibliography of  
the writings of Jonathan Swift. 1963. p.  
192-198)
```

Figure 27i
Bibliographical Record

```
001 $a 402123456 [Bodleian internal  
number]  
004 $a OCLC 11392229  
500 $a Bookplate of the Benedictine  
Monastery in Grafschaft; inscribed by  
Charles Dickens  
700 10 $a n 78087607 $4 ins  
710 20 $a n 87654321 $4 fmo  
852 81 $a n 80020306 $b BOD $c Vet. A3  
d.9999
```

Figure 27j
Copy Record

allowing massive amounts of data to be exchanged within a couple of seconds. The idea of accessing and retrieving large amounts of data on a regular, frequent, and perhaps online basis is now realizable.

The development of client-server architecture and the Z39.50 standard allow the possibility that the data can be retrieved on demand, so that a bibliographical record may be viewed in which some elements are generated locally and some remotely, with the remotely generated elements (say the current authorized LC form of name, held at LC or OCLC) interfacing seamlessly and transparently with the local elements. The new 856 field for the Universal Resource Locator can hold an address for a remote resource which can be accessed directly from the record. If this can be done, then online links to remote authority files can be made.

There will always be, however, local records not immediately reflected in a national or supranational database. Even if all cataloging institutions eventually

have online read and write access to such a database, there may be local needs which it would be wrong to impose upon a shared record (for example, a local reference in a general work). An essential structural requirement for any authority record in the future must be a place to hold local as well as national record-identifying numbers. Matching algorithms would have to be run regularly on local-only authority records to establish whether that identity had appeared in the wider authority file. The existence of a uniform-title authority file linked to a name authority file should make such algorithms much more feasible, as the potential matches could be cross-checked between files. Recent proposals have indeed been made which would use name-title algorithms to correct and update authority files.

Conclusion

I am proposing major revisions of MARC in order to allow records to maximize the benefits of computer databases, on the one hand, and high-speed data networks, on the other. This involves a fundamental shift away from the AACR2 philosophy of description of, plus access to, physical items. It also involves a re-analysis based on a modeling of the methods by which "works" (in the abstract) are created and made available, relying on the attributes of such works and the processes which they may undergo.

Today, the present system is held down by the weight of its own history. The fate of the MARBI discussion paper on uniform titles (*Possible Changes . . .* 1994) makes the point as clearly as possible: we must suffer poor analysis and unhelpful records because too much has been invested in the structure we use. Michael Gorman has suggested that we will "remain in the iron grip of an outdated technology expressed in the MARC system" (Gorman 1992, 93). The only way out of this impasse is to make a bold leap. I have suggested such a leap, but at the same time have tried to indicate the ways in which the changes could be made without having to throw away the cumulated achievement of the last three decades. We must sidestep the problems by building around them instead of trying to break through them head-on.

My proposed revision produces a three-tiered structure consisting of the abstract work, the publication, and the copy. The current MARC bibliographical record embraces all three of these, but at its core is the publication. Current MARC authority files could form the basis for structures for the abstract work. Formats for the copy are not well developed: only the US MARC holdings format begins to address the problem, and

there is a need for major enhancement if that is to serve as a basis for the copy model.

In such a structure, bibliographical description is by-and-large an insignificant part of the publication model, and the major access to information is by the "abstract work." The major thrust of cataloging should therefore be in building conceptual models of works and their agents, by which people usually try to find such works. Much of the descriptive element of AACR2 could simply wither on the vine. It is easy to forget that, in principle, MARC is independent of AACR2 and of any other set of cataloging rules. In some cases this approach also frees us from AACR2's Procrustean approach to nonbook materials. To give just one example: an obvious attribute of every map is that it depicts a place, yet AACR2's sole concession to this fundamental attribute is to instruct catalogers to add a place name to the title area of the description.

Throughout this discussion I have focused, for convenience, on written works and published texts and have used the terms "Text model" and "Publication model." In a more generalized formulation, it may be more appropriate to use the terms that are now coming to be widely used, and to call them the "Work model" and the "Manifestation model."

The logical conclusion to this train of development is that, as far as library catalogs are concerned, AACR2 is not so much a plateau of maturity as an evolutionary blind alley. It is a code of the age of the catalog card and printed bibliography. While such purposes and products still exist, AACR2 still has a place; however, for computerized library catalogs and in the world of the virtual library, it is time to retreat to first principles and initiate the development of AACR3 as an object-oriented cataloging code based on the principles outlined above.

Acknowledgments

Pat Oddy of the British Library very kindly read and made many helpful comments on an early draft of this paper. Frank Sadowski was very helpful in guiding me to and supplying me with the *Guidelines for Bibliographic Description of Reproductions*. Bob Strauss kindly allowed me to cite his contribution to the AUTOCAT Discussion Group.

References

- American Library Association. 1992. Committee on Cataloging. Description and Access Multiple Versions Taskforce. "Guidelines for bibliographic description of re-

- productions," (report presented to CC:DA, ALCTS, ALA).
- Bierbaum, Esther Green. 1994. "A modest proposal: No more main entry." *American Libraries* 25, no. 1 (Jan.): 81-84.
- Bodleian Library. 1986. *The first printed catalogue of the Bodleian Library, 1605: A facsimile. [Catalogus librorum bibliothecae publicae quam . . . Thomas Bodleius eques auratus in academia Oxoniensi nuper Instituit.]* Oxford: Clarendon Press.
- "Considerations on a patron record format." 1993. MARBI discussion paper, D.P. 71 (28 May).
- Gorman, Michael. 1979. "Cataloging and the new technologies." In *The nature and future of the catalog: Proceedings of the ALA's Information Science and Automation Division's 1975 and 1977 institutes on the catalog*, ed. by Maurice J. Freedman and S. Michael Malinconico, 127-36. Mansell: Oryx.
- . 1992. "After AACR2R: The future of the Anglo-American Cataloguing Rules." In *Origins, content, and future of AACR2 Revised*, ed. Richard P. Smiraglia, 89-118. ALTCS Papers on Library Technical Services and Collections, no. 2. Chicago and London: ALA.
- Leazer, Gregory H. 1992. "An examination of data elements for bibliographic description: Toward a conceptual schema for the USMARC formats." *Library resources and technical services* 36, no. 2 (Apr.): 189-208.
- Lunt, Colin. 1993. *Technical limitations of the MARC format: A discussion document*. Version 1.0. Users of Book Industry Standards, Working Group on Bibliographic Formats (1 Feb., 1993).
- "Possible changes to treatment of uniform titles in the USMARC bibliographic and authority formats." 1994. MARBI discussion paper, D.P. 72 (22 Feb.).
- Strauss, Bob. 1993. "The future of cataloging (and libraries)." *AUTOCAT: Library cataloging and authorities discussion group* (2 Dec.).
- Tillett, Barbara B. 1989. "Bibliographic structures: the evolution of catalog entries, references, and tracings." In *The conceptual foundation of descriptive cataloging*, ed. Elaine Svenonius, 149-65. San Diego, Calif.: Academic Press.
- Wainwright, Eric. 1991. "Implications of the dynamic record for the future of cataloguing." *Cataloguing Australia* 17, no. 3/4 (Sept./Dec.): 7-20.

THE CAT SHAKESPEARE WHO KNEW

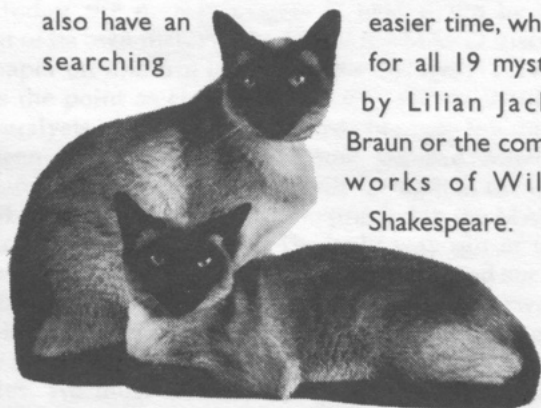
In this popular feline felony by Lilian Jackson Braun, the detective team of Jim Qwilleran, Koko and Yum Yum solve the mystery of Senior Goodwinter's death. Did the death of the mild-mannered typesetter revolve around adultery, gambling, drugs or espionage? Was he the fourth victim to suffer the Pickax Picayune jinx?

Koko provides clues to the mystery by pushing volumes of Shakespeare off shelves in the family library for Qwill to read.

It is only when Koko dislodges a copy of Hamlet that the crime is finally solved.

With WLN's CD-ROM LaserCat, Koko, Yum Yum and Qwill would have had a much easier time locating the Shakespearean clues. You'll also have an easier time, whether

for all 19 mysteries by Lilian Jackson Braun or the complete works of William Shakespeare.



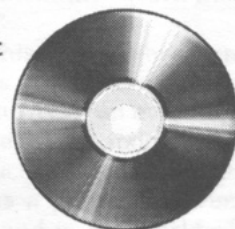
LASERCAT

LaserCat's 4 million bibliographic records provide a solution to every librarian's cataloging needs twenty-

four hours a day, seven days a week.

LaserCat is designed to meet the diverse requirements of large and small collections, whether it's the Pickax Public Library, a university consortium

in Minneapolis or your own unique library. Of course, LaserCat is available in Macintosh, Windows and DOS versions.



So, if you don't have an expert detective like Koko on your staff, you should consider adding another kind of cat - LaserCat from WLN.

For a free demo diskette, call 1-800-DIAL WLN. Don't ask for Koko, but ask for LaserCat from WLN.

WLN

Turning Information Into Knowledge

What would collection developers do in the absence of local collections? Collection developers partition the universe of documents, privileging selected documents for acquisition and retention. Their selections of documents, based on expected demand and perceived value, direct the actions of technical services staff. A comparable selective privileging of documents, based on demand and value, and a similar direction of a reoriented technical services activity appears needed with networked electronic resources. Collection developers will be needed for value-based privileging more than for demand-based decisions.

What would collection developers do if and when the emerging environment of networked electronic resources were to lead to the absence or reduced significance of local collections?

There can be little doubt that the rise of network-accessible electronic resources appears to offer a substantial alternative in access to library resources to both of the traditional options of holding materials in local collections or of dependence on interlibrary loan. As one example, J. P. Migne's *Patrologiae cursus completus* (Latin series), a 221-volume compilation of the nineteenth-century writings in Latin of the early Christian fathers, has become available in electronic form. Thus, it would appear that almost any material of interest to library users could become available in electronic form. Consequently, access to networked electronic resources offers, in principle, an alternative to the assembling of local collections—although there remain many significant problems to be addressed.¹ Indeed, in the special case of indexing and abstracting services, the searching of remote databases has already tended to replace subscription to locally-held paper editions.

We are not and need not be concerned here with whether, how, how completely, or under what circumstances access to networked resources will replace local collections in paper. Instead, we examine the consequences that such a replacement might have for collection developers.

The mere idea that access to networked electronic documents could substantially replace local collections raises intriguing questions about the purposes of local collections and the work of collection developers. What do collection developers do now for library users? At the superficial level of process, what they do is clear: They make a stream of decisions telling the library's technical services staff which items should be acquired and cataloged for the local collection, which should be weeded,

and, the largest category, what should not be acquired. However, the *purpose*, as distinguished from the *process*, is less clear. There are large bodies of literature on collection developing, on how new library-related technologies are evolving, and on the issue of a shift in emphasis from ownership to access.² Important though these writings are, they are primarily concerned with process (with "how") and so can distract attention from examination of purpose (of "why"). A period of time in which significant technological change is occurring provides a good opportunity not only to emphasize purpose, but also to theorize about the nature of what we do.

There has, in recent years, been a significant shift away from viewing the role of technical services staff as that of managing a grand apparatus for establishing usable local collections, toward a notion of their role as providing the users' gateway to the bibliographic universe.³ In what follows, I argue for and predict a comparable change in the role of collection development staff.

The Purposes of Collections

Library collections can be viewed in terms of four roles.⁴

1. Preservation: If a document is to be available now and in the future, then at least one copy needs to be collected somewhere;
2. Dispensing: The principal reason why libraries acquire most of the materials that they add to their local collections is not for preservation but because of the need to provide convenient physical access to materials where and when users want them;
3. Advisory (or bibliographic): The array of materials on the shelves can itself alert the reader to what is available, just as any bibliography or catalog can. Certainly the array on the shelves is a selective, incomplete guide. It is limited to what has been added to that collection and, within that, is biased towards the less-popular material that happens to be on the shelves at any particular time. Nonetheless, a library collection plays an advisory role like that of a selective bibliography, drawing attention to material that has been identified as worth adding to the collection. Browsing books has some attraction over browsing in bibliographies or among catalog records. It is largely for this advi-

Michael Buckland is Professor at the School of Information Management and Systems, University of California, Berkeley. His e-mail address is buckland@otlet.berkeley.edu.

sory role that the materials are arranged by a subject classification scheme.

4. Symbolic: It can also be argued that the acquisition of extensive collections and, especially, rare and prestigious materials adds status to a library and its parent institution.

Of these four roles, the vast preponderance of actual expenditure is attributable to the second and third roles—the two roles that seem most likely to be transformed by the new technology.

Book Selection Theory

Lionel McColvin's classic *The Theory of Book Selection for Public Libraries* (1925) begins with the following words: "Book selection is the first task of librarianship. It precedes all other processes—cataloguing, classification, or administration—and it is the most important. No matter how thorough and efficient the rest of the work may be, the ultimate value of a library depends upon the way in which the stock has been selected."⁵

McColvin begins with the need to have available the materials that will be in demand, not only because of the practical need to have in stock what users ask for, but also because there will be no benefit derived from acquisitions that are not used. No demand means no use and, therefore, no benefit, McColvin argues. But he also argues that a passive adapting of collections to demand would betray the mission of the library. "If, however, we consider the library as a social force with the power to direct to some extent man's demand, (or, to use the usual expression, if we consider the library as an *educational* force) we will not be content to leave demand our only consideration."⁶In brief, McColvin, like other writers on collection development, provides a frame with two dimensions—demand and value—which can be regarded as related to the dispensing and advisory roles of collections noted above.

There is a specialized literature by Lister, Slote, Trueswell, and others on the quantitative techniques that can be used to adapt collections to patterns of demand.⁷ However, much discussion of collection development—especially that concerning "balanced" collections—has to do with librarians' value-laden departures from purely demand-driven collection development. It is a matter of deliberately privileging some documents over others as a separate matter from the anticipation of demand. The books placed in the shelves—in the reader's face, so to speak—carry an implicit endorsement: these you should know about; these are good books for you; or, these are books you will like. Other

materials, those not selected for (or weeded from) the collection, are actively, though implicitly, treated as less suitable for readers. They have been deliberately judged less satisfactory for whatever reason. The evaluations may well be derivative, based, for example, on book reviews or faculty recommendations.

Collection developers will consider both demand and value. Nevertheless, it is important to recognize that the value aspect, the privileging of some works over others on grounds of value, is different in kind from the narrowly logistic task of anticipating demand to maintain high immediate availability. The specialized literature on quantitative collection development techniques developed by Lister and others for ensuring high immediate availability is essentially irrelevant to this second task of the privileging of some documents over others.

Networked Electronic Resources

With materials on paper, the development of well-selected local collections dominates the quality of service, as McColvin stressed. There are gradations of local availability (on reserve, out on loan, in storage, etc.), but there is a basic, binary distinction: What is held locally is accessible and what is not held locally is relatively inaccessible. Recourse to interlibrary borrowing is an unsatisfying substitute for local holdings. Put simply, by the process of selection the collection developer is imposing a partitioning of the universe of library materials into two populations: those that are to be made more accessible by being held locally and those that are to be kept less accessible because they are not placed in the local collection. Structurally, the effect is the same as the compilation of a selective bibliography or the result of an online search: Some items are selectively brought forward to the reader's attention—acquired, listed, and retrieved, respectively; while others—those not acquired, listed, or retrieved—still exist and are still accessible, but are left in less-accessible obscurity.

With networked, electronic resources, however, this binary distinction between local (and accessible) and nonlocal (and less conveniently accessible) becomes less clear. In principle, all resources become equally (more or less) accessible. The distinction between locally held and not locally held loses significance and, in a sense, all collections everywhere can become locally accessible collections. A consequence relevant to technical services is that the local catalog, essentially a guide to what is locally held, loses its past pre-eminence relative to union catalogs, remote catalogs, bibliographies, and guides to network-accessible resources.

Local Collections in the Electronic Network Environment

"In the library of the future there will be no library. All will be navigation!" I have heard from an academic acquaintance. This declaration does not imply an end to library service or to librarians (who are all good navigators, of course). Some of us may prefer "bibliography" to "navigation" as the term of choice and, obviously, some repositories (collections) of electronic materials must be located somewhere. What is placed in question, though, is the future of the *local* collections that have hitherto dominated library service and have accounted for most of libraries' expenditures once the full costs of selecting, acquiring, processing, and housing local collections are properly attributed.

Clearly there will continue to be local collections of two kinds: (1) materials on paper, microfilm, and other localized media for which the reader and the document must be in the same place; and (2) "caches" of electronic documents that are used often enough to justify local storage.

The population of documents in the localized caches of electronic documents will be transient and transparent. In general, users need neither know nor care which documents are stored locally and which are not at any given time. Automatic algorithms (essentially the same as those developed by Lister and the others noted above) can be designed, based on expected frequency of use, unit storage costs, and the costs and delays ("latency") of obtaining a copy from remote storage, to adjust the cache dynamically. The design of such algorithms is a task for industrial engineers rather than subject specialists. Site licenses will need to be negotiated, but that seems likely to become more like negotiating a blanket order than traditional title-by-title, copy-by-copy book selection.

There are localized electronic media, notably CD-ROMs, that must be selected and acquired like books or microfilms. But CD-ROMs can be put on networks and their contents can be stored in repositories. CD-ROMs seem a transitional technology or a temporary storage device.

As for the local collections in localized media (paper, microform, etc.), indications are that they will be a diminishing portion of what is used and will no longer define, as they had in the past, a library's ability to provide service. So the question remains: What will collection developers do as local collections diminish in significance relative to networked electronic resources? Will collection developers' professional lives be enriched by the assignment of other, different duties? The answer, I suggest, lies in the closer examination of the

purposes of what collection developers do now, or in McColvin's distinction between demand and value.

In traditional collection development, or in building local collections of localized media, the single act of acquisition is the common response to both demand and value. One procedure addresses both concerns. One cannot know by examining a book on the shelf whether its acquisition resulted from an expectation of demand, from some ascribed value, or from some combination of the two. In an electronic, networked environment, however, considerations of demand and of value diverge because they require different courses of action. The logistics of catering to high demand, in detail a matter of caches, can be delegated as primarily a mechanical task, and one would have a welcome opportunity to economize on collection developers if that were all they were to do.

But what of the concern for *value* in the electronic library environment, for the privileging of some books over others? If there is more to collection development than responding to demand, then the value-laden role of privileging some resources over others must be continued unless the purpose of library service is to change. How is this other, remaining task of collection developers to be accomplished? If it can no longer be accomplished obscurely, combined with the logistical task of meeting demand by placing copies on the shelves, it will need to be addressed directly and separately.

Collection developers and technical services staff have been more closely co-conspirators in achieving the library's mission than libraries' organizational charts have indicated. Acquisitions departments and catalogers process only those items that the collection developers select. Collection developers and technical services staff play complementary and interdependent roles in establishing an ordering of the universe of documents, or in determining the relative accessibility of different documents, for their local users. There is no clear reason why that purpose and that partnership should cease. Quite the reverse, as electronic resources multiply, the need for a convenient ordering of differentiated accessibility increases. Technical services have been evolving away from being the grand apparatus that constructs the local collection and towards the provision of gateways to the bibliographic universe. Value judgments are still needed concerning which resources are most suitable for any given user group.

The privileging of the better and, by default, the nonprivileging of the rest remains an important needed service. If we accept McColvin's statement for library service, written in relation to documents printed on paper, on what grounds would we deny it for library service using documents on disk drives? Should the choice of technology for the storage medium determine the mission of the library? Let us re-read McColvin in

relation to networked electronic resources: "If, however, we consider the library as a social force with the power to direct to some extent man's demand, (or, to use the usual expression, if we consider the library as an *educational* force) we will not be content to leave demand our only consideration."

What collection developers do now is to select items for local acquisition. The purpose of that process is to manipulate the universe of documents so that some subsets—on grounds of demand and/or value—are made more accessible than others. Some documents are made visually prominent by being placed on the local shelves, while others are deliberately not. As paper documents on shelves cease to be the technological medium of choice, different but comparable procedures are needed for a different technological medium.

In the provision of navigational (alias bibliographic) tools to the universe of electronic documents, it seems contrary to common sense and to the historic traditions of library service to make all material equally accessible. There is simply too much material and, while we may agree that it ought all to be accessible, it would be unhelpful to make it all equally accessible. Some degree of differentiation is needed. Some items are demanded more frequently than others, some may be regarded as more valuable than others. It has not been the purview of technical services staff to select which items should be privileged over others, but rather to implement that privileging for (and only for) the documents designated by the collection developers.

The design of a gopher service provides a simple example. Not all items can be equally accessible at the same, highest hierarchical level. It makes for efficiency if items that will be looked for often are given a privileged place high in the hierarchy of gopher levels. It makes for effectiveness if items we regard as valuable for library users and wish to have seen are placed high in the hierarchy. Other items remain accessible, but can be left to be in (or to be presented as if in) deeper, less convenient levels of storage. If book selection is seen as deciding which items to privilege, then the need for those with that ability would appear to increase as local paper collections diminish relative to networked electronic collections—and the traditional partnership with technical services should be just as tight.⁸

Privileging Networked Resources

What collection developers will do procedurally in the future with the new technology can be expected to differ in various ways from what was done in the past with the old technology:

1. Hitherto the privileging of documents has been dominated by a binary division: items acquired for the local collection and those not acquired or not retained. In the environment of networked resources such an abrupt division seems improbable. A much finer gradation of degrees of accessibility and privileging seems likely.
2. Hitherto all users of a given library have been supplied with one and the same collection. This "one-collection-for-all" approach has been technologically inevitable, but it is Procrustean rather than democratic or egalitarian, since different users have different needs and users are unlikely to be equally well served by what a single collection contains or by the way it is arranged. The popularity of branch and departmental libraries arises from their being customized to special needs as well as from their geographical convenience. With the new technology, different forms of access (multiple "clients") can be designed for different interest groups within the local population served.
3. Because of the inherent localness of local collections, collection development work has been specific to each location and has resulted in massive geographical inequalities in library holdings. Library users with similar interests but located at different sites have received radically different service. With the new technology it may well be that the task can and will become more specific to topical areas than to locality, which opens new opportunities for cooperative efforts. Similar forms of access could be shared by those with similar interests but who are at different locations—those in a topical rather than a geographical community.
4. Access and ownership, always separate in principle but hitherto rarely separate in library practice, are expected to diverge in the electronic network environment. One consequence is that "collection development" is much less limited to one's own materials, or that one can "privilege" the materials of others. A second consequence is that users can also do the same, designing ways of privileging materials not belonging to them, thereby providing self-help alternatives to supplement or replace the work of professional collection developers.
5. Because new technology is significantly more flexible than that of paper and cardboard, making multiple alternative approaches more feasible, issues of value and privileging now need to be addressed in more complex ways. Fundamentally, this involves a shift from traditional standardized

provision in one or a few ways towards more flexible systems, designed to be adaptive and more responsive to users' desires to invoke their own preferences in exploring the universe of documents. (This trend is already evident in discussions of online catalog design.⁹)

6. The notion of "materials budget" will evolve. Historically, a component of the cost of making privileged documents more accessible, a different deployment in approach to this budget is inevitable if the traditional purpose of library service is to be sustained in a changed environment.
7. Because the evaluative, privileging role will become a separate task from catering to demand, it can be expected to become a performance with greater visibility and accountability, with interesting consequences for professional behavior and new mechanisms for monitoring the quality of professional performance—as has already happened for catalogers.¹⁰
8. This new need for the evaluative skills of collection developers arises from the emergence of the new environment of networked resources, not from changes in the roles of local collections.

Conclusion

What collection developers will do depends on how one views what they do now. At the superficial, procedural level, it seems that there will be a much-reduced need for collection developers. However, if we are to take seriously the purposes underlying the procedures used in the development of local collections, then even though new and different technology brings new and different procedures, the fundamental purposes and the expertise needed for selection (as opposed to acquisition) will remain crucial in providing value judgments to guide those who design network access mechanisms. Thus, in the emerging electronic environment, whatever the fate of local collections, we should expect a fundamental continuity in what collection developers will do.

References

1. For an excellent analysis of these problems see Clifford A. Lynch, *Accessibility and Integrity of Networked Information Collections* (Washington, D.C.: United States Congress. Office of Technology Assessment, 1994; BC-TCT-109).
2. For example, Joel S. Rutstein, Anna DeMiller, and Elizabeth Fuseler, "Ownership Versus Access: Shifting Perspectives for Libraries," *Advances in Librarianship* 17 (1993): 33–60.
3. Barbara A. Norgard, Michael G. Berger, Michael Buckland, and Christian Plaunt, "The Online Catalog: From Technical Services to Access Service," *Advances in Librarianship* 17 (1993): 111–48; Michael Buckland, "From Catalog to Selecting Aid," *ALCTS Newsletter* 5, no. 2 (1994): Insert A–D. ("From Catalog to Gateway: Briefings from the CFFC," 2).
4. Michael K. Buckland, "The Roles of Collections and the Scope of Collection Development," *Journal of Documentation* 45 (1989): 213–26.
5. Lionel McColvin, *The Theory of Book Selection in Public Libraries* (London: Grafton, 1925), 9. For a convenient summary of McColvin see Robert N. Broadus, *Selecting Materials for Libraries* (New York: H. W. Wilson, 1973), chapter 2; for a broader historical background see Rinaldo Lunati, *Book Selection: Principles and Procedures* (Metuchen, N.J.: Scarecrow Press, 1975).
6. McColvin, *Theory*, 37. Emphasis in original.
7. For example, Winston C. Lister, *Least Cost Decision Rules for the Selection of Library Materials for Compact Storage*. Ph.D. dissertation, Purdue University, West Lafayette, Indiana, 1967 (ERIC ED 027 916); Richard W. Trueswell, "User Satisfaction vs. Size of Holdings at Three Academic Libraries," *College and Research Libraries* 30 (1969): 204–13; Michael K. Buckland, *Book Availability and the Library User* (New York: Pergamon, 1975); Stanley J. Slote, *Weeding Library Collections* (Littleton, Colo.: Libraries Unlimited, 1975); John A. Urqhart and N. C. Urqhart, *Relegation and Stock Control in Libraries* (Stocksfield, United Kingdom: Oriel Press, 1976); Gary S. Lawrence, "A Cost Model for Storage and Weeding Problems," *College and Research Libraries* 42 (1981): 139–47; R. E. Staynor and V. E. Richardson, *The Cost-Effectiveness of Alternative Book Storage Programs* (Clayton, Victoria, Australia: Monash University, School of Librarianship, 1983).
8. Lynch, *Accessibility*, 78–81.
9. For example, Michael K. Buckland, Mark H. Butler, Barbara A. Norgard, and Christian Plaunt. "OASIS: A Front-end for Prototyping Catalog Enhancements," *Library Hi Tech* 40 (1992): 7–22.
10. Ruth A. Hafter, *Academic Librarians and Cataloging Networks: Visibility, Quality, Control, and Professional Status* (Westport, Conn.: Greenwood, 1986).

A New Service on the Information Superhighway

Authority ExpressSM

If you have been searching for an easy way to authority control your library's current cataloging, try LTI's **Authority Express** service.

With **Authority Express**, a library uses the Internet to transmit a file of newly cataloged bibliographic records to LTI (via FTP). Overnight, LTI processes the records through its state-of-the-art authority control system. Then, at the library's convenience, it logs into LTI's FTP server to retrieve fully authorized catalog records, along with linked LC name and subject authority records.

Authority Express

- Keeps authority control current at an affordable price
- Integrates easily into existing workflows
- Lowers cost by reducing staff time spent on catalog maintenance
- Provides next-day turn around for up to 5,000 catalog records
- Accepts records for processing even if LTI did not perform the original authority control

"Authority Control for the 21st Century"



LIBRARY TECHNOLOGIES, INC.

1142E Bradfield Road Abington, PA 19001
(215) 576-6983 Fax: (215) 576-0137
(800) 795-9504 email: LTI@LibraryTech.Com

Precision and Recall in Title Keyword Searches

Monica Cahill McJunkin

This study examines whether adjacency operators in title keyword searches are effective in improving precision. Title keywords from a random sample of titles in economics were searched in FirstSearch with and without adjacency of the keywords specified. Subject headings of the items retrieved were compared with the sample title subject headings to determine the degree of match or relevancy, and figures for precision and recall were calculated. When keywords were discipline-specific, adjacency operators improved precision with little degradation in recall. Systems that allow positional operators or rank output by proximity of terms may increase search success.

Online catalogs provide many opportunities for creative subject access, including keyword searches. While keyword searches in controlled vocabulary fields allow access to subject headings when entry terms or word order are not known, titles also contain subject-rich terms. These keywords use the author's own terminology, which is often more current than the *Library of Congress Subject Headings (LCSH)*,¹ and can be combined or related to each other in order to vary the search. This study investigates the extent to which title keywords convey subject content and compares the relative effectiveness of searching title keywords via two different strategies.

Unlike searches in non-keyword-based systems, which must match the beginning of the field, keyword searches involve identifying the requested terms at any position in the field being searched. Multiple terms can be combined in a search using the Boolean operators AND, OR, and NOT. Word stems or truncated terms, as well as positional operators, can be specified. These operators can specify the order in which the terms appear, their proximity to each other, or that the terms be adjacent to one another. The options in keyword searching allow the user to broaden or narrow a search as needed.

Peters and Kurth determined from a study of dial-access transaction logs at the University of Missouri-Kansas City that library patrons were using title keyword searches as a form of uncontrolled vocabulary search.² In other studies, users were observed using title terms for subject access both in the catalog and while browsing the shelf.³ These studies make a case for the existing use of subject access through title keywords, but show no evidence of the success of these searches, nor

the relative success of different types of keyword searches.

Other studies have found title terms used for subject searching: Larson has described the decline of subject searching and the concomitant rise in title keyword searching over a six-year period,⁴ and Ensor describes several studies which show a rise in keyword searching of all types.⁵ Both authors note that keyword use rises with catalog experience. Connell observed that experienced users perform title keyword searches as a lead-in to the controlled vocabulary,⁶ and Peters and Kurth recommend this method in addition to using title keyword searches alone.⁷

When users perform title keyword searches as a subject approach to the catalog, how good are the results? More specifically, do items that contain the same terms in their titles cover the same topic, and are certain title keyword search strategies more effective than others for subject searching?

Literature Review

Characteristics of Title Keyword Searches

Title keyword searching has some advantages over controlled vocabulary subject access. Titles terms are more likely to agree with the user's terminology and serve as a complement to the assigned subject headings⁸ and have been found by Jamieson to overlap very little with subject cross-references.⁹ Bates found that subject experts in economics consistently preferred headings that were more precise than the subject headings assigned to works, and that they particularly disliked the subheading "economic conditions" because of the variety of meanings covered by it.¹⁰ In her dissertation (described by Connell¹¹) Bates also found that users had particular difficulty with subject heading matching for economics items; economics headings tend to be complex, often including subheadings for time periods and geographic regions.

However, title keywords are only as good as the author makes them. Even after articles, prepositions, and conjunctions are removed from consideration, generic terms like "report" remain, as do metaphors and cute, catchy phrases. Often, synonyms and spelling variations compound the problem. In general, keyword searching exhibits a lack of tolerance for misspellings

Monica Cahill McJunkin is Technical Services Librarian at Otterbein College, Westerville, Ohio.

and variations of any kind.¹² Lastly, because the terms are taken out of context, keywords that are retrieved may be used in a different manner than was intended by the searcher.¹³

Evaluation Methodologies

Many studies have attempted to evaluate the usefulness of title keyword searching. Connell used keywords from abstracts in *Book Review Digest* to determine to what extent book descriptions match terms in subject headings or titles.¹⁴ She also looked at fields that are not commonly used, such as the subtitle or other title information, to determine their potential in retrieving items. All words in the descriptions were considered keywords except the following: a, an, and, at, by, for, from, how, in, of, on, the, to, with. Connell compared all keywords from the abstracts with fields in the bibliographic record and found that, for books for which no match was found between the description and the subject headings or LCSH cross-references, 27.8% matched title keywords. Of the books remaining, over a third produced matches in the subtitle field. While some of these last matches were with terms that indicate the form of the item, subtitles often provide meaningful keywords when the title proper contains a catchy phrase. This study indicates that titles and subtitles may be useful for subject access; however, the percentage of matches reflects only those titles for which subject headings and subject cross-references failed to produce a match.

In a study that took the opposite approach, Gerhan compared the usefulness of terms in titles and subject headings by determining whether they were likely to be used by patrons desiring items on that topic.¹⁵ Catalog cards were examined for terms which had a reasonable probability of being search terms; these terms were subjectively rated according to whether Gerhan thought that patrons would use them for subject access. He found that title keywords are effective retrieval terms about 55% of the time (including 10% of the instances in which subject headings are absent or extremely lacking), but that subject headings were effective about 85% of the time, and therefore made a better first choice for searching. Gerhan concluded that terms from subject headings and titles are often complementary, and that use of both methods may be the most productive.

Cherry took yet another approach. While Connell started with book descriptions and Gerhan started with catalog entries in order to determine the likelihood that books would be found based on keywords in the bibliographic record, Cherry examined unsuccessful subject searches (defined as those with zero hits).¹⁶ Actual user subject searches were converted to subject keyword searches, title searches, title keyword searches, and sub-

ject cross-reference searches. Title keywords were the most useful, retrieving records in 62% of the cases, as opposed to subject keywords and subject cross-references, which were each successful 33% of the time. Title searches, which must match the beginning of the title, were successful at retrieving records 43% of the time. Although these searches were only performed with requests that had already failed with traditional subject access, this study does indicate that title keyword searching is a useful addition to subject searching, especially since the search terms employed were actual patron search requests.

Aanonson compared some retrieval sets from subject searches he performed while evaluating keyword searching on six university catalogs.¹⁷ He determined that title keyword searches not only retrieved useful items not found with subject keyword searches, but that they provide useful starting points for getting into the controlled vocabulary. He also found that additional useful records were retrieved when the series title was included in the title keyword search as well.

Evaluation Measures

The previous studies cited did not evaluate the relevance of retrieved items. Books were not examined to determine content and search terms were accepted as accurate portrayals of desired subjects. The number of records retrieved was the main consideration. However, large retrieval sets can be a disadvantage while searching if the user must browse through many records looking for useful items. Larson attributes a decline in subject searching over a six-year period to increases in database size and resulting user frustration with large retrieval sets.¹⁸ He notes that keyword-based systems are more likely to cause information overload for the user, and favors ranking of output records according to the number of search terms contained in each record. Yee, on the other hand, suggests that keyword indexing may be improved by locational data to allow searching of keywords combined into "phrases"¹⁹; while Lancaster and his coauthors include the limiting of keyword searches by date, language, or other factors as a way of improving subject access.²⁰

Evaluating retrieved records according to their relevance can be a complex issue. First, one must distinguish between pertinence and relevance. Relevance has been defined as a "relationship between a document and a request" and pertinence as the "relationship between a document and an information need."²¹ In other words, a relevant item is one that matches the search request, while a pertinent item is one that is judged useful by the user. In the absence of real users with actual information needs, the relevance of an item can be agreed upon by a

group of subject experts. Kemp views relevance as objective and pertinence as subjective, drawing a parallel in psychology with denotation of words (objective) and connotation (subjective).²² Others disagree, claiming that, whenever relevance decisions are made by individuals or groups of individuals, they must be subjective and dependent upon a variety of external factors. In either case, making relevance decisions based on a subjective measure of topicality can be appropriate for initial evaluations of a system's retrieval capabilities.²³

When no users are involved and items are not available for evaluation, other methods must be used to determine relevance. Although finding a matching LC subject heading does not guarantee search success, Bates claims that a matching score between search terms and subject headings is a good measure of success.²⁴ The LC subject heading should provide one best heading, controlling for synonyms and related terms, and should match the scope of the item.

Once a method for determining relevance has been chosen, records can be weighted according to their usefulness. Unlike known item searches, subject searches need a "measure of degree of success"²⁵—some items are more relevant than others. In a study of database coverage for periodical indexes, Sharma weighted items according to the following scale:²⁶

Fully relevant	1.00
Half or moderately relevant	0.50
Marginally relevant	0.25
Irrelevant	0.00

This type of weighting procedure could apply to any method of deciding relevance.

Once relevance is determined, recall and precision figures put the relevance figures into perspective. Recall is defined as the percentage of relevant documents retrieved; precision is defined as the percentage of retrieved documents that are relevant.²⁷ Generally, recall and precision are inversely related; for example, term truncation improves recall but degrades precision. While precision is easy to calculate because it is based on the ratio of relevant items retrieved to total items retrieved, recall is harder to estimate because it involves the ratio of relevant items retrieved to total relevant items in the database, which is impossible to know. Lancaster has suggested estimating the total number of relevant items by having several users perform parallel searches (i.e., use different search strategies), then combining the total number of relevant items retrieved to represent the number of relevant items in the database.²⁸ While this will not disclose indexing failures in the database, it can highlight the usefulness of different search strategies.

Summary

To date, research on title keyword searches has typically focused on comparisons of title keyword searches with subject or subject keyword searches. Book descriptions, user searches, or "made up" terms served as the source of keywords, and title keyword searches were found to complement subject searches. In general, title keyword searching is often characterized by poor precision because the terms are taken out of context. Truncation and word stemming can increase recall, but searches in large databases often suffer more from a lack of precision. Better precision may be obtained by the use of word proximity or adjacency operators, which combine keywords into meaningful phrases.²⁹ However, it is not known to what extent this degrades recall.

Objectives and Definitions

The objectives of this study are:

- To determine the levels of precision and recall obtained with title keyword searching for titles in economics;
- To determine the levels of precision and recall obtained with title keyword searching for titles in economics modified by adjacency operators to create keyword "phrases"; and
- To compare the levels of precision and recall obtained via the two methods in order to determine which is the more effective means of subject access.

Unlike previous studies, titles are the source of keywords and provide the searched fields. In effect, works on the same topic are assumed to use the same title terms if they are to be used for subject access. Because of the difficulties with subject access that have been described above, economics was chosen as the subject field for this study. Keywords include all terms except the stop words used by Connell: a, an, and, at, by, for, from, how, in, of, on, the, to, with.³⁰ The number of keywords varies from search to search: Keywords were searched singly or in combination with stopwords in the title delineating the search groups. For example, the two keyword groups for the title *Low-Income Housing in the Developing World* are "low-income housing" and "developing world." When more than one term is included, a Boolean AND is implicit in the search.

Searches were performed on *FirstSearch*, using the WorldCat database, which is equivalent in coverage to the OCLC Online Union Catalog. *FirstSearch* was chosen to provide the largest possible coverage with the least

bias introduced by individual institutional holdings. Title keyword searches cover the title proper field, as well as other title information, uniform titles, added titles, and series titles.

Two different search strategies were used: In the first, keyword(s) were entered and searches were performed without regard to word order or proximity. In the second, keyword(s) were entered with adjacency operators that specified the exact phrases to be matched.

Relevance was determined by the degree of LC subject heading match between the source title and the retrieved title. While not ideal, this provides an objective measure that can be used for other studies and separates the issue of the adequacy of the indexing language from the comparison of keyword search strategies. Sharma's weighting scale has been adapted for this study:

Exact subject match	1.00
Broader or narrower	0.50
Related	0.25
No match	0.00

Broader and narrower matches include headings that omit or include, respectively, subdivisions, in addition to those defined by the *LCSH* hierarchy. Similarly, related matches include headings with the same main heading but different subdivisions. Because all subject headings from source and retrieved titles were considered, it was possible for an item to receive a relevance score greater than 1.0.

The denominator for calculating recall, the total number of relevant documents in the database, was estimated using the union of the unique relevant records (weighted score) retrieved via the two methods and the number of unique records obtained via an exact phrase subject heading search, using headings from the source titles. Recall, then, is the number of unique relevant records (weighted score) retrieved, divided by this denominator. Precision is simpler and is defined as the number of unique relevant records (weighted score) retrieved, divided by the total unique records retrieved.

It is important to note that the scope of this study does not involve comparing title keyword searching with subject searching; the objective is to examine the effect of adjacency operators on recall and precision for title keywords only. While it is not possible to know how many relevant items may be missed by title keyword searches, the extent to which titles containing the same terminology are on the same subject is an important consideration in title keyword searching. In this respect, it is the relative effectiveness of two title keyword search strategies that is being examined.

Methodology

Precision and recall of title keyword searches in economics were obtained by analyzing search results from the *FirstSearch WorldCat* database.

Sample

The members of the target population were monograph titles in economics; the accessible population sampled was made up of the titles in *Economics and Business*, an annotated bibliography that was published from 1984 through 1986. The entries are numbered, which facilitates sampling, and they cover a range of subtopics on economics, such as monetary theory, international economics, and industrial organization, so that the vocabulary is varied. In addition, because the titles are from a limited number of years, the searches could be limited to these years. The vast majority of titles fall between 1983 and 1985, so only titles in that range are included in the sample.

A random sample of titles was drawn from the bibliography using a table of random numbers. The sample size was chosen to obtain 90% confidence with a margin of error of ten percentage points, using a standard formula. After the results of a pilot study suggested that the standard deviation for precision and recall was estimated to be 0.5, a sample size of 68 was chosen.

Procedures

Before title keyword searching began, data for the sample titles were collected. First, each title was searched in the OCLC Online Union Catalog in order to make a list of the LC subject headings assigned to each title. It is important to stress that no subject headings from other authority lists were considered in this study. For this reason, this search could not take place on *FirstSearch* because the source of subject headings is not displayed with the records. Second, the keyword combinations to be searched for each title were recorded and numbered.

Title keyword searches in the *FirstSearch WorldCat* database began after first limiting the searches by language (English) and year of publication (1983–1985). In order to simplify and standardize the searches, keywords were not searched in various forms—such as truncation, word stemming, elimination of plurals, or various spellings. In order to keep the retrieval sets manageable, further limits were imposed: If any subject heading search for a title yielded more than one thousand records, all searches for that title were limited to one year. If any title keyword search yielded more than five hundred records, one hundred records were system-

atically sampled from the retrieved set. Large retrieval sets that fell into one of the following categories were not sampled and were instead omitted from the study:

- Keyword contains the bibliographic format of the item (guide, directory);
- Keyword contains the presentation or treatment of data (analysis, survey);
- Keyword contains a generic geographic or chronological term (area, nation, era); or
- Keyword contains a broad geographic or chronological term (United States, twentieth century)

When in doubt, the retrieved records were sampled. These limits were necessary because of the significant proportion of overly large retrieval sets: Out of 350 possible searches, 85 (or 24%) of the retrieval sets contained 500 or more records. Of these, 66 (19%) were sampled and 19 (5%) matched one of the categories described above and were omitted. Appendix A gives a breakdown of searches by retrieval set size.

Each title was searched using both strategies. When a keyword stood alone in a title, both strategies were completed in one search. For example, in the title *Agricultural Development in Bangladesh*, "Bangladesh" stands apart from the other keywords, so adjacency operators can not be used. The syntax of the search statements for this title were:

```
s ti:agricultural development
s ti:bangladesh
s ti:agricultural w development
```

In the first search, the terms "agricultural" and "development" could appear in the retrieved records in any combination of searched fields in either order. In the third search, the terms must appear in the same field together as a phrase in the order specified. The second search retrieves the keyword "bangladesh" for both methods.

Transaction logs from the search sessions were downloaded. The source title, which should appear in the retrieval sets, was removed from consideration. Then data for each search were recorded, including: the search number and lists of retrieved records (by OCLC record number) in columns for exact match, broader, narrower, and related. The relevance scores for each retrieved record were determined by comparing subject headings of the sample titles with those of the retrieved titles using the tenth edition of the *Library of Congress Subject Headings*. This edition was chosen because it most closely corresponds with the time period in which the sample titles would have been cataloged.

The relevance of each retrieved record was determined by rating each of the retrieved item's subject

headings as an exact match (1.0), broader heading (0.5), narrower heading (0.5), related heading (0.25), or no match (0.0).

In order to estimate a denominator for recall, subject headings from the sample titles were searched. Only exact subject heading matches were to be included; however, the WorldCat subject headings can not be searched exactly. Exact phrase searching is available on subject heading fields, but the various segments of the subject headings are indexed separately. Thus, a search for "Government lending — United States," which is stated as "sh=(government lending and united states)," will also retrieve "Government lending — Law and legislation—United States," "Government lending—United States — Handbooks, manuals, etc.," as well as a record with the pair of headings "United States — Small Business Administration" and "Government lending — Arkansas." Each retrieved set of records was edited to remove the extraneous headings.

Data Analysis

Figures for recall and precision were estimated for each search method by the following method: For each search, the total number of relevant records was calculated. Then recall and precision for each search were estimated for each of the two keyword search strategies using the following formulae:

$$R = \frac{r}{k + s - l_{ks}}$$

$$P = \frac{r}{t}$$

where

- R = recall
- r = number of relevant records retrieved in this search
- k = number of relevant records from title keyword searches for this keyword grouping (w/o adjacency)
- s = number of records from exact subject heading searches for this title
- l_{ks} = number of records contained in both k and s (overlap)
- P = precision
- t = total records retrieved in this search

The set of relevant records retrieved when adjacency is specified is always a subset of the set of relevant records retrieved when adjacency is not specified; therefore, only the larger set is necessary for calculating the denominator for recall. Precision and recall were then averaged for each title.

Because every subject heading in a retrieved record is evaluated for relevancy, an individual record may have a relevancy score greater than 1.0; thus, precision for a search (and average precision for a title) may also be greater than 1.0. Also, since the denominator for recall includes each relevant record only once, but a retrieved record may have a relevancy score greater than 1.0, it is possible for recall for a search to be greater than 1.0.

Discussion

Out of 68 titles, 29 required no sampling of retrieved records and 39 contained retrieval sets which were sampled due to their size. These are referred to as "non-sampled titles" and "sampled titles," respectively.

Subject Heading Searches

The inability to search for exact subject heading matches was unexpected. The searches for three of the titles were limited to one year because the retrieval sets for individual subject heading searches were greater than one thousand. For title 47, searching "sh=population" limited to 1984, retrieved 1291 records—only 40 of which were found to contain the exact subject heading "Population." Subject headings containing subdivisions pose an additional problem: For title 54, searching "sh=(small business and united states)" retrieved 836 records—only 235 of which contained the exact heading "Small business—United States." Not only were other subdivisions also included in the retrieved records, but "Small business" and "United States" did not have to appear in the same heading in order for a record to be retrieved. While the flexibility allowed by this system has some advantages, ranking of output according to the degree of match to the search statement should be incorporated. If all records containing "Population" were listed before variations including subdivisions, evaluation of records would have been easier and the search would not have had to be limited to one year.

Unusual Relevancy Scores

Some titles have no precision or recall scores, while others have scores exceeding 1.0. Undefined scores occur when searches retrieve no records (other than the sample title). When there is no set of retrieved records, calculating precision is impossible and calculating recall, although theoretically possible if exact subject heading searches retrieved a nonzero set, is meaningless. There are no undefined precision and recall scores in the sampled titles (because there was always at least a

sample of one hundred records retrieved); undefined scores occur in four of the non-sampled titles. For example, title 13 is "Socio-economic accounting." Searching for these keywords either with or without adjacency operators retrieves no records other than the sample title.

High precision and recall scores occurred for both sampled and non-sampled titles. As described in the data analysis section, retrieved records may receive a relevancy score greater than 1.0. This usually occurred when there were few exact subject heading matches and title keyword searches retrieved small sets of records consisting mostly of other editions of the same work. These cases have been included in the overall data calculations, even though they are artificially large. The alternative would be to evaluate all retrieved records in order to eliminate those which are considered duplicates of the sample, or of each other, to determine that the retrieval sets contain only unique records. With so many records, however, and none of the items in hand, this alternative is not feasible. It was assumed that these duplicates are evenly distributed throughout the retrieval sets and would not affect the comparison between the two search strategies.

Non-sampled Titles

Table 1 contains summary data for the non-sampled titles. For convenience, title keyword searches performed without adjacency specified are referred to as keyword searches, and title keyword searches performed with adjacency specified are referred to as phrase searches. Confidence intervals were generated

Table 1
Non-sampled Titles

Type of Score	Confidence Interval
Keyword precision	.3016 ≤ .4402 ≤ .5787
Phrase precision	.3580 ≤ .5335 ≤ .7090
Difference between strategies (keyword precision – phrase precision)	-.1488 ≤ -.0783 ≤ -.007
Keyword recall	.1134 ≤ .1745 ≤ .2357
Phrase recall	.0933 ≤ .1488 ≤ .2043
Difference between strategies (keyword recall – phrase recall)	.0043 ≤ .0323 ≤ .0604

using the z-statistic at a level of significance of 0.10. The mean precision scores for keyword and phrase searching are 44% and 53%, respectively. The confidence intervals overlap quite a bit, yet it is clear that higher precision was obtained from phrase searching. The mean difference between the scores is 7.8%.

The mean recall scores for keyword and phrase searching are 17.5% and 15%, respectively. The loss in recall obtained with phrase searching is much less than the gain in precision, and the confidence intervals almost totally overlap. Recall scores for keyword searches are, on average, 3% higher than recall scores for phrase searches.

The values of the keyword and phrase scores relative to each other are what would be expected; phrase searching results in higher precision with only a slight loss in recall. In other words, the number of unwanted records (in which the keywords were used in a different context than desired) that were eliminated exceeded the relevant records that were missed. It is significant to note that because keywords occurring singly were searched singly for both strategies, the difference in precision is not as large as it might be if only multiple-word keyword phrases were included in the study. They were included to obtain a more realistic sense of how the strategies would perform against each other in natural settings, in which keywords would often be searched singly despite user strategy preferences or system defaults. The precision and recall obtained when single keywords are excluded is explored later.

Sampled Titles

Table 2 contains summary data for precision and recall for the sampled titles. Confidence intervals were again

Table 2
Sampled Titles

Type of Score	Confidence Interval
Keyword precision	.1376 ≤ .2352 ≤ .3328
Phrase precision	.1520 ≤ .2541 ≤ .3562
Difference between strategies (keyword precision – phrase precision)	-.0334 ≤ -.0189 ≤ .0045
Keyword recall	.1353 ≤ .2153 ≤ .2954
Phrase recall	.1214 ≤ .2020 ≤ .2827
Difference between strategies (keyword recall – phrase recall)	-.0014 ≤ .0133 ≤ .2796

generated using the z-statistic. The mean precision scores for keyword and phrase searching are almost identical—23.5% and 25.4%, respectively, with a mean difference of less than 2%. The confidence intervals overlap almost completely. Mean recall figures are also similar to each other—21.5% and 20% for keyword and phrase searching, respectively. The lack of difference between strategies may be due to the preponderance of single keyword searches in the sampled titles, for which keyword and phrase searching are identical.

Comparing the data in table 2 with the data for non-sampled titles in table 1, it is apparent that some factor is causing a significant difference in the relevance scores. Precision for sampled titles is much lower than the precision for non-sampled titles. The keywords in the sampled titles are more likely to be general, nondiscipline-specific terms (hence the need to sample from large retrieval sets). These terms are used in a variety of contexts, resulting in a significant number of nonrelevant records. In addition, the sampled titles tend to contain more keywords and more keyword groupings than the non-sampled titles. This reduces the probability that any one keyword (or keyword grouping) adequately describes the content of the item and lessens the probability that retrieved records will have matching subject headings.

Number of Keywords or Keyword Groups

An analysis of the titles that contained three or more single keywords to be searched shows that 13 out of a total of 15 are sampled titles. Confidence intervals for all subsequent tables were generated using the t-statistic for a two-tailed test at the 0.10 level of significance. As summary data for the 13 sampled titles in

Table 3
Sampled Titles with Three or More Single Keywords

Type of Score	Confidence Interval
Keyword precision	.0438 ≤ .1829 ≤ .3220
Phrase precision	.0401 ≤ .2017 ≤ .3633
Difference between strategies (keyword precision – phrase precision)	-.0447 ≤ -.0188 ≤ .0071
Keyword recall	.1220 ≤ .2114 ≤ .3009
Phrase recall	.1235 ≤ .2154 ≤ .3072
Difference between strategies (keyword recall – phrase recall)	-.0165 ≤ -.0039 ≤ .0086

table 3 shows, precision for both keyword and phrase searching is low—only 18% and 20%, respectively—which indicates that searching single keywords, which are less specific in meaning than multiple word phrases, lowers precision.

Recall is very similar to the recall obtained for all of the sampled titles: 21% and 21.5% (table 3) versus 21.5% and 20% (table 2). This is slightly higher than the recall obtained for the non-sampled titles (see table 1). Although the general terms found in the sampled titles result in larger retrieval sets containing many nonrelevant records, they pick up more of the relevant records.

Titles that contained four or more keyword groups were combined in a like analysis. While this set does not quite overlap completely with the titles containing three or more single keywords, it yielded similar results, but with slightly lower recall. Scores for both precision and recall may be lower than for the entire group because, as the number of keyword groups increases, it is less likely that any one group approximates the content adequately. Fewer relevant records are retrieved and thus recall suffers as well as precision.

Conversely, 11 of the non-sampled titles contained only a single keyword or keyword group. The figures for precision are much higher than those of the entire non-sampled group—75% (keyword searches) and 94% (phrase searches). However, these figures are artificially inflated by the occurrence of small retrieval sets that include records that are duplicates or close matches to the sample title and thus have unusually large relevance scores.

Number of Subject Headings

Since relevance is evaluated based on subject heading matches, the number of subject headings assigned to the sample titles was analyzed to see if this affected precision and recall.

Summary data for titles that have three or more subject headings assigned to them are shown in table 4. For both non-sampled and sampled titles, there is little difference in precision and recall due to strategy. For non-sampled titles, precision and recall both dropped significantly from the scores for all the non-sampled titles (see table 1.) This may indicate that titles with three or more subject headings have complex or varied topics that cannot be described with only one or two subject headings. Non-sampled titles tend to have fewer, more specific keyword groupings (2.25 per title versus 4 for sampled titles) and may have specific terms which match none of the subject headings. For sampled titles, recall is not significantly different from the recall obtained for all of the sampled titles. Precision, however,

Table 4
Titles with Three or More Subject Headings

Type of Score	Confidence Interval
<i>Non-sampled Titles</i>	
Keyword precision	.1285 ≤ .2425 ≤ .3565
Phrase precision	.1207 ≤ .2567 ≤ .3927
Difference between strategies (keyword precision – phrase precision)	-.0674 ≤ -.0142 ≤ .0390
Keyword recall	.0343 ≤ .0920 ≤ .1496
Phrase recall	.0240 ≤ .0769 ≤ .1298
Difference between strategies (keyword recall – phrase recall)	.0039 ≤ .0151 ≤ .0262
<i>Sampled Titles</i>	
Keyword precision	.1258 ≤ .3306 ≤ .5353
Phrase precision	.1398 ≤ .3482 ≤ .5566
Difference between strategies (keyword precision – phrase precision)	-.0441 ≤ -.0176 ≤ .0089
Keyword recall	.0629 ≤ .2343 ≤ .4058
Phrase recall	.0321 ≤ .2042 ≤ .3762
Difference between strategies (keyword recall – phrase recall)	-.0015 ≤ .0302 ≤ .0619

is improved (33% and 35% versus 23.5% and 25.4%; see table 2). This is likely because sampled titles tend to contain more keyword groupings, which have a greater chance of match against several subject headings.

Confidence intervals are displayed in table 5 for titles that have only one subject heading. For the non-sampled titles, precision for both strategies is similar to precision for all non-sampled titles (46% and 50% versus 44% and 53%, in table 1). Recall, however, is greatly improved. Since non-sampled titles have, on average, fewer keyword groupings than the sampled titles, when only one subject heading is assigned, the topic of the work is covered by one phrase. Thus, recall may be improved because the few keyword groupings are more likely to match the single subject heading.

For the sampled titles, precision and recall are both lower than for sampled titles as a whole. This is probably

Table 5
Titles with One Subject Heading

Type of Score	Confidence Interval
<i>Non-sampled Titles</i>	
Keyword precision	.2804 ≤ .4560 ≤ .6396
Phrase precision	.2921 ≤ .5023 ≤ .7124
Difference between strategies (keyword precision – phrase precision)	-.0951 ≤ -.0423 ≤ .0106
Keyword recall	.1377 ≤ .2764 ≤ .4152
Phrase recall	.1226 ≤ .2521 ≤ .3816
Difference between strategies (keyword recall – phrase recall)	-.0190 ≤ .0243 ≤ .0677
<i>Sampled Titles</i>	
Keyword precision	.0505 ≤ .0918 ≤ .1331
Phrase precision	.0460 ≤ .1144 ≤ .1827
Difference between strategies (keyword precision – phrase precision)	-.0540 ≤ -.0226 ≤ .0088
Keyword recall	.0695 ≤ .1504 ≤ .2314
Phrase recall	.0589 ≤ .1454 ≤ .2319
Difference between strategies (keyword recall – phrase recall)	-.0143 ≤ .0005 ≤ .0244

because the large number of terms or phrases do not match well, individually, to a single subject heading.

Single Keywords Excluded

Lastly, keywords that were searched singly were removed from consideration in order to determine their effect on precision and recall. Table 6 shows the data for these titles. Only seven titles remained in the "sampled" titles category. This demonstrates that single keywords tended to be general terms that resulted in large retrieval sets and thus required sampling.

For the non-sampled titles, confidence intervals were generated using the z-statistic. Precision levels for keyword and phrase searching are 49.5% and 59%, respectively, with a mean difference of 7%. These are similar to but slightly higher than the levels for non-sampled

Table 6
Single Keywords Excluded

Type of Score	Confidence Interval
<i>Non-sampled Titles</i>	
Keyword precision	.3513 ≤ .4956 ≤ .6399
Phrase precision	.4277 ≤ .5928 ≤ .7578
Difference between strategies (keyword precision – phrase precision)	-.1137 ≤ -.0692 ≤ -.0247
Keyword recall	.1211 ≤ .2255 ≤ .3299
Phrase recall	.0837 ≤ .1912 ≤ .2987
Difference between strategies (keyword recall – phrase recall)	.0209 ≤ .0497 ≤ .0786
<i>Sampled Titles</i>	
Keyword precision	.1258 ≤ .3306 ≤ .5353
Phrase precision	.1398 ≤ .3482 ≤ .5566
Difference between strategies (keyword precision – phrase precision)	-.0441 ≤ -.0176 ≤ .0089
Keyword recall	.0629 ≤ .2343 ≤ .4058
Phrase recall	.0321 ≤ .2042 ≤ .3762
Difference between strategies (keyword recall – phrase recall)	-.0015 ≤ .0302 ≤ .0619

titles including single keywords (shown in table 1) and much higher than the levels obtained for sampled titles including single keywords (shown in table 2). As expected, the removal of the single keywords, which are more general in meaning, results in higher precision with a significant difference between keyword and phrase searching strategies. Recall levels are 22.5% and 19%, with a mean difference of 5%.

Confidence intervals were generated for the sampled titles using the t-statistic with 6 degrees of freedom. Precision is low, only 20.3% and 21.6% for keyword and phrase searching, respectively. Even though the single keywords have been excluded, these titles still contain general terms that required sampling of retrieval sets; the use of nonspecific terms results in lower precision with little difference between search strategies. Recall, at 21% and 22%, is not significantly different from the recall obtained with non-sampled titles.

Summary and Conclusions

This study has examined precision and recall obtained from title keyword searches performed with and without adjacency operators. When keywords are limited in meaning, precision is significantly improved by the use of adjacency operators, and recall declines to a lesser extent. Due to the design of this study, other factors—such as the level of specificity of the terms, the length of the sample title, and the number of subject headings assigned to the sample title—were larger influences.

Overall, precision and recall were quite low. Many exact subject heading matches were missed by title keyword searches. Precision can be improved by choosing search terms carefully; discipline-specific, subject-rich terms are best. Care should also be taken when using title keyword searches as a lead-in to the controlled vocabulary. Users should be aware of the standard terminology in the field and the level of specificity needed. As with any search, users who are not familiar with a subject's terminology may not end up with the single best heading. For example, a keyword search for "macroeconomics" would pull up records with the subject heading "Macroeconomics." However, the user may really have something like "supply-side economics" in mind, but does not know how to phrase it for a search. Users who are not subject experts should consult the *LCSH* or online cross-references in order to find the correct terminology.

On the other hand, users should also be knowledgeable about the online system in order to use it effectively. A user who does not know that *FirstSearch* may retrieve terms from several fields in the same record may be confused by the results. A search for "industrial structure," for instance, may retrieve a record with "pricing structure" in the title and "industrial commission" in the series title. Here, certainly, knowledge about the system's search logic and the availability of adjacency operators is helpful. Although the results of this study seem to support the use of adjacency operators to improve searching effectiveness, a user for whom absolute recall is more important may wish to use a broader search strategy.

Title keyword searching, with or without adjacency operators, is available in many online catalogs and is sure to be added to more in the future. While evidence suggests that library patrons are using title keyword searching as a means of subject access, we have few measures of its effectiveness and, as database sizes increase, precision will be an ever-growing problem. Whenever title fields are searched, alone or in combination with other content-bearing fields such as subject headings or notes, precision requires that title terms be indicative of the content of the item.

Additional research is needed to clarify the extent to which adjacency operators affect precision and recall. Future research could repeat this study with a larger sample size, using only discipline-specific and/or multiple-word keyword phrases in order to magnify the relationship between adjacency operators and precision and recall. Other disciplines could be examined or the focus could be on journal article titles. Future research could take another direction and repeat this study using truncation of terms or proximity operators in place of adjacency.

If studies support certain strategies as being more helpful than others, this could have implications in several areas. First, more systems can be designed to support these strategies. Second, users can be instructed on the relative merits of different strategies, either formally or through help screens. Third, retrieval systems could be designed to default to certain strategies under some conditions or to rank the output based on adjacency or proximity, in order to increase search success without increasing user effort. If *FirstSearch* ranked output of keyword searches by the proximity of the terms to each other in the record, users may not have to learn these strategies at all. It is hoped that the results of this study will help build a framework in which to view all keyword search strategies.

References

1. Lois Mai Chan and Theodora Hodges, "Subject Cataloguing and Classification: The Late 1980s and Beyond," in *Technical Services Today and Tomorrow*, comp. Michael Gorman (Englewood, Colo.: Libraries Unlimited, 1990), 74–85.
2. Thomas A. Peters and Martin Kurth, "Controlled and Uncontrolled Vocabulary Subject Searching in an Academic Library Online Catalog," *Information Technology and Libraries* 10 (Sept. 1991): 201–11.
3. Michelene Hancock, "Subject Searching Behaviour at the Library Catalogue and at the Shelves: Implications for Online Interactive Catalogues," *Journal of Documentation* 43, no. 4 (1987): 303–21; and Michelene Hancock-Beaulieu, "Evaluating the Impact of an Online Library Catalogue on Subject Searching Behaviour at the Catalogue and at the Shelves," *Journal of Documentation* 46, no. 4 (1990): 318–38.
4. Ray R. Larson, "The Decline of Subject Searching: Long-term Trends and Patterns of Index Use in an Online Catalog," *Journal of the American Society for Information Science* 42, no. 3 (1991): 197–215.
5. Pat Ensor, "User Characteristics of Keyword Searching in an OPAC," *College & Research Libraries* 53 (Jan. 1992): 72–80.
6. Tschera Harkness Connell, "Librarian Subject Searching in Online Catalogs: an Exploratory Study of Knowledge Used" (Ph.D. diss., University of Illinois at Urbana-Champaign, 1991).
7. Peters and Kurth, "Controlled and Uncontrolled Vocabulary Subject Searching."

8. Rao Aluri, D. Alasdair Kemp, and John T. Boll, *Subject Analysis in Online Catalogs* (Englewood, Colo.: Libraries Unlimited, 1991).

9. Martha M. Yee, "System Design and Cataloging Meet the User: User Interfaces to Online Public Access Catalogs," *Journal of the American Society for Information Science* 42, no. 2 (1991): 78-98.

10. Marcia J. Bates, "Factors Affecting Subject Catalog Search Success," *Journal of the American Society for Information Science* 28 (May 1977): 161-69.

11. Connell, "Librarian Subject Searching in Online Catalogs."

12. John Akeroyd, "Information Seeking in Online Catalogs," *Journal of Documentation* 46, no. 1 (1990): 33-52.

13. John N. Olsgaard and John Edward Evans, "Improving Keyword Indexing," *Journal of the American Society for Information Science* 32 (1981): 71-72.

14. Tschera Harkness Connell, "Techniques to Improve Subject Retrieval in Online Catalogs: Flexible Access to Elements in the Bibliographic Record," *Information Technology and Libraries* 10 (June 1991): 87-98.

15. David R. Gerhan, "LCSH *in Vivo*: Subject Searching Performance and Strategy in the OPAC Era," *Journal of Academic Librarianship* 15, no. 2 (1989): 83-89.

16. Joan M. Cherry, "Improving Subject Access in OPACs: An Exploratory Study of Conversion of Users' Queries," *Journal of Academic Librarianship* 18, no. 2 (1992): 95-99.

17. John Aanonson, "A Comparison of Keyword Subject Searching on Six British University OPACs," *Online Review* 11, no. 5 (1987): 303-13.

18. Larson, "The Decline of Subject Searching."

19. Yee, "System Design and Cataloging Meet the User."

20. F. Wilfred Lancaster and others, "Identifying Barriers to Effective Subject Access in Library Catalogs," *Library Resources and Technical Services* 35, no. 4 (1991): 377-91.

21. F. Wilfred Lancaster, *Information Retrieval Systems: Characteristics, Testing and Evaluation*, 2d ed. (New York: Wiley, 1979).

22. D. A. Kemp, "Relevance, Pertinence and Information System Development," *Information Storage and Retrieval* 10 (1974): 37-47.

23. William Hersh, "Relevance and Retrieval Evaluation: Perspectives from Medicine," *Journal of the American Society for Information Science* 45, no. 3 (1994): 201-6.

24. Bates, "Factors Affecting Subject Catalog Search Success."

25. Lancaster and others, "Identifying Barriers to Effective Subject Access."

26. V. S. Sharma, "Information Resources for Recreation, Leisure and Sport," *Database* 5 (August 1982): 32-50.

27. Aluri, Kemp, and Boll, *Subject Analysis in Online Catalogs*.

28. Lancaster, *Information Retrieval Systems*.

29. Chan and Hodges, "Subject Cataloguing and Classification."

30. Connell, "Techniques to Improve Subject Retrieval in Online Catalogs."

Appendix A: Number of Records Retrieved

No. of Records Retrieved	No. of Searches
0-9	134
10-49	66
50-99	21
100-499	44
500-999	18
1,000-4,999	33
5,000-9,999	17
10,000+	17



LAUNCH INTO A NEW ERA: Resources for the Future Librarian from ALA Editions

Academic Libraries as High-Tech Gateways: A Guide to Design and Space Decisions

Richard J. Bazillion and Connie Braun

Finally, in one complete and readable resource, librarians find out what it takes to make academic libraries gateways to vast resources of electronic and printed information. The authors demystify the planning process to help librarians make design and space decisions for online information access, electronic study centers, and other automated systems. Practical topics include attributes of a 21st-century library; creating the "intelligent library"; furnishing, equipping, and testing the building; and, libraries as teaching instruments. If your library belongs to another century, you can't afford to be without this book.

\$40.00pbk. • 225p. • 1995 • ALA Order Code 0656-7-0011

The Internet Troubleshooter: Help for the Logged-On and Lost

Nancy Regina John and Edward J. Valauskas

This compact question-and-answer book is a lifeline for Internet surfers lost at sea. The authors have taught hundreds of Internet users at every level and understand the common problems that arise in the midst of online activity. An index, section dividers, screen illustrations, glossary, and real-life examples make solutions easy to find, understand, and enact. With its spiral binding, the guide is perfect for use by your terminal.

\$25.00spiral-bound • 145p. • 1994 • ALA Order Code 0633-8-0011

Future Libraries: Dreams, Madness and Reality

Walt Crawford and Michael Gorman

Amidst the rush to all-electronic libraries, here's one book that may save your library. This bestseller presents an authoritative counter-argument to the over-hyped virtual library, submits evidence for the continuing value of print collections among diverse media, explains how to deal with the enemies of today's libraries, furnishes a "survival guide to the serials crisis," and exposes the fantasies associated with the electronic-publishing industry.

\$25.00pbk. • 198p. • 1995 • ALA Order Code 0647-8-0011

FROM THE LIBRARY AND INFORMATION TECHNOLOGY ASSOCIATION

Internet Connections: A Librarian's Guide to Dial-Up Access and Use, 2nd Edition

LITA Monographs, #6

Mary E. Engle et al.

Includes a snapshot of the networking services available at time of publication, a discussion of the concepts and terms, and a bibliography of the notable guides that document search techniques, navigational tools, and available information resources. Provides a general overview of electronic service providers offering dial-up access to the Internet, along with pointers to other valuable sources of information.

\$24.00pbk. • 280p. • 1995 • ALA Order Code 7793-6-0011



155 N. Wacker • Chicago, IL 60606

To order, call 800-545-2433 and press 7



COMMUNICATIONS

"Two Thumbs Up!" Library Applications for Video Technology

Daren Callahan

The medium of video is ubiquitous, one with which we are all familiar. It is a predominant means of communication today, and videotapes with diverse applications flood the market. Producers extol the unique features and flexibility of video technology and recommend its use for educational, promotional, and motivational purposes. In Morris Library at Southern Illinois University at Carbondale we have successfully incorporated this very contemporary tool into the library environment. A small group of staff members participated in a project designed to produce various library-oriented videos. The project began with a series of workshops in basic video production. One of the first programs was a fifteen-minute videotape on the Open Court Collection, the archives of a noted Illinois publishing company and one of the largest collections in the library's Manuscripts Department. The potential for the use of video in libraries is extraordinary. In order to capitalize on this potential, a video producer has now been employed full-time to continue the program within the library and to assist video production throughout the campus community.

Background

The video project, originally designed by the dean of Library Affairs, was intended to produce a series of videotapes for use in the library's different divisions. The videotapes would be used to update patrons about new services in the library, for faculty presentations,

and for promotional purposes. We retained complete control of the presentations from beginning to end, primarily by utilizing resources within the library itself. Staff members from various departments were recruited to develop the subject matter and direct each production. While none of the recruits had prior experience with video production, the technical expertise and video editing equipment necessary to ensure a professional, effective product were available in the library's Instructional Services Division, a unit whose activities include course design and evaluation, implementation of instructional technology, and production of audio and videotapes for the campus community.

A half-dozen production workshops, lasting approximately one hour each, were conducted by the head of the library's Instructional Development Division and were held every few weeks. During these sessions five novice directors learned about the principles and steps involved in creating a successful video with the maximum impact. A good video, we were told, is effective because it is lively, interesting, and concise. In order to maintain optimum interest and attention, the videos were designed to last no more than ten to fifteen minutes. Once the purpose and proposed audience for each were clearly defined, writing the scripts and compiling the shot lists followed.

In Special Collections, an early suggestion to prepare a training film for student workers on filing in the card catalog was abandoned and replaced with a more interesting proposal to present material from the Open Court Collection that could be shown to researchers, students, and anyone interested in but unfamiliar with this material. The collection includes not only the archives of the Open Court Publishing Company, but also the personal papers of its editor, Paul Carus. Hundreds of

boxes house an extensive correspondence with scientists, philosophers, and writers from around the world. There are copies of all the books published by the company, as well as galleys, page proofs, and many rare pamphlets and journals from the latter part of the nineteenth century and early years of the twentieth century. In addition there are many photographs, paintings, drawings, and illustrations—a wealth of material that could be incorporated into the presentation.

Since the Open Court Collection consists chiefly of books, galleys, manuscripts, and correspondence—material that typically lacks inherent visual interest—video presentation of such material offered a significant challenge. By interspersing examples from the illustrative material and correspondence when appropriate, we were able to relieve the monotony of the "talking head" syndrome that often dominates straight narration and create a kind of documentary of the collection. We used photographs of many of the well-known correspondents who had visited Carus through the years. Some of their letters contained lively passages and amusing closing remarks, which we were able to use. There were also many intriguing early photographs from Tibet, Nepal, and Sri Lanka. We also included photographs that were taken as Special Collections staff members packed up the materials, which had been stored in an old four-story mansion in LaSalle, Illinois, the original home of the press. The staff members appeared smiling in dust masks as they worked through a century's accumulation of papers, books, and coal dust in the many rooms throughout the house.

The Actor

Generally, most people become nervous in front of a camera because they lack professional acting experi-

Daren Callahan is Special Collections Cataloger at Southern Illinois University, Carbondale.

ence. Without sufficient preparation, they may be apprehensive about their performance, their appearance, or their delivery; consequently, they may doubt their ability to be convincing. Everyone in the department suddenly became shy when initial discussions for the video began. No one wanted to appear in the film.

Rather than have a member of the Special Collections staff talk about the material in the video, we asked a current researcher to discuss the contents and significance of the collection. Our researcher, a historian and teacher, was preparing a book on Paul Carus and the Open Court Publishing Company and had visited the library many times. He had studied the letters, books, and pamphlets in the collection and was willing to speak about a subject that was close to his heart. He spoke extemporaneously and enthusiastically about material with which he was familiar and, because he was an experienced communicator, he appeared quite relaxed in the video.

We were fortunate to have at hand someone with an excellent presentation style who had a good rapport with the camera. If an experienced speaker is unavailable, it is wise to consult a media coach for advice. With video, it is crucial to present a message that the audience can relate to; thus, the speaker must appear positive and credible. The performance must be polished and engaging; otherwise, one should plan to get enough raw footage so that the piece can be carefully edited to convey the desired message. Video-makers can also choose to use voice-over exclusively. While this method eliminates the need for a person to appear on camera, the stills and narration used in the video must be sufficiently interesting, and the camera movement varied enough, to maintain the attention of the audience.

Editing the Videotape

Coordinated by the staff in the library's Instructional Services Division, the production had one graduate assistant from that department—an experienced videographer—who was committed to the production from beginning to end. After filming was complete, we worked closely together in assembling the video, virtually frame by frame. The final raw footage compiled for what eventually became a fifteen-minute videotape amounted to almost four hours. This included the section we filmed with the researcher, complete with varying camera angles, "filler" footage (i.e., looking through folders, reading letters, working at his laptop computer), and numerous stills of materials from the collection. Working from the time code on the videotape, we extracted the most useful commentary and the appropriate stills to illustrate the narrative. This master tape then had to be reworked into a smooth, rational sequence of beginning, middle, and end. A scripted introduction employing voice-over was incorporated to present the viewer with the subject matter and its context.

All of this demanded a great deal of hard work that was complex, time-consuming, and prone to a certain amount of trial and error. It was also fascinating and enjoyable. The editing of the Open Court videotape was completed over a five month-period during half-day sessions arranged when schedules permitted, at which times we explored the capabilities of the editing console. An experienced editor familiar with the system can edit a fifteen-minute videotape, including artistic transitions and graphics, in less than a week.

The library owns a variety of camcorders in 1/2-inch, 3/4-inch, and super-8 format. It also houses a complete video production studio

with an integrated A/B roll operating system, including a switcher, 8-channel stereo audio board, preview and playback monitors, CD player, and time-base correctors. This sophisticated system enabled us to mix two tape sources and to vary the presentation of images with high-end digital picture manipulation. When editing, we were able to incorporate transition effects such as dissolves, wipes, sweeps, freeze-frame, and fragmented screens—all of which are attention-getting and provide visual variety. Such effects also help to deal with unexpected gestures, fleeting expressions that cross a face, nuances in speech, or anything distracting that must be edited out of the final version.

One important lesson learned from this early production was that a video presentation is most successful if one person directs the program and has the final authority for aesthetic and technical decisions. "Video-by-committee" should be avoided. If a piece is required to satisfy an assortment of directors with diverse opinions, the program becomes muddled, lacks a unifying design, and pleases no one in the end.

Production Costs

We were most fortunate to have the complete support of the library administration in this endeavor. This support allowed a faculty member some release time from regular duties during a twelve-month period to participate in the workshops and to complete the videotape. We also had unlimited access to state-of-the-art editing equipment, which helped to guarantee an impressive presentation. In addition, because the equipment and staff were available within the library itself, there were no external costs.

A complete video editing unit with a full range of capabilities can

cost up to \$85,000. One less-expensive alternative is an Amiga product, the Video Toaster, commonly known as "video in a box." This software allows one to edit and manipulate images, create different backgrounds, and incorporate graphics. An integrated system utilizing the Video Toaster is very efficient, capable of producing a professional-looking videotape, and costs about \$13,000. A good camcorder can cost about \$1,500, and an S-VHS (Super VHS) product is recommended. In addition, an experienced operator is essential—someone who knows the equipment and understands shot composition, lighting, and all aspects of editing.

If video production expertise and equipment are not available in-house, libraries can usually contract with a video production company. The current industry standard is approximately \$1,000 per finished minute of videotape, but this estimate will vary locally depending on crew costs, locations, complexity of editing, and style of the production. While videotape production may appear to involve an expensive investment, if videotapes can be used on multiple occasions for training or for promotional purposes—in fundraising perhaps—then it can be considered a sound investment. If contracting with an outside agency, it is important to make efficient use of time, to develop the script in close communication with the videographer, and to learn the jargon of those responsible for the editing to ensure that you get a program that meets the institution's needs. With a clearly defined objective, it is possible to get a reliable cost estimate for the production.

Some academic libraries can contact the video production unit within their institution—such as the film, television, or broadcasting departments—where one can find enthusiastic faculty and students with

good experience and access to equipment who are often looking for projects. If an institution owns a video camera, lights, and sound recording equipment, it is possible to do the taping oneself and then contract with a video production company or an independent videographer to do the editing. This should cost between \$130–\$250 per hour depending upon your requirements.

Library Applications

Video is indeed a flexible tool suited to a wide range of applications. Libraries are currently embracing an assortment of new technologies and can only benefit by exploring the many uses of videotape. The Open Court video has been well received by the surviving Carus family members and the board that currently administers the company. It reassures the donors of the value of their gift and underscores the library's commitment to providing access to these materials. In the sense that video exemplifies modern technology, innovative programs produced in the library would likely win approval from forward-looking administrators. Moreover, a stylish videotape enhances the library's image by drawing attention to the importance of the services we provide.

In addition to the Open Court production, another videotape was made, titled "The SIUC Library: The Future Is Now," which features the many technological developments recently incorporated in the library, along with plans for new electronic tools. This video has been shown locally to university administrators, state officials, the Friends of the Library, and important visitors, in addition to being presented at a number of professional conferences. Another videotape in progress is intended for students and explains the use of the Reserves/Self-Instruction

Division. The library's next video production will focus on a program to implement distance learning capabilities within the library. When necessary, portions of each of these videotapes can be easily updated by the staff in Instructional Services with the addition of new footage and/or commentary.

With the Open Court project we discovered that presenting collections on video introduces a new, altered view of familiar library holdings—the books and paper that fill our stacks. Seeing materials in a new context with an information mix—narrative combined with images—prompts new interest and appreciation and may induce a reevaluation of these materials. One can elect to give a quick-paced tour of the contents of a collection in the form of a video montage or go into a deeper analysis of the history and significance of the material with a video-essay. Video could also be an exciting element in a Special Collections exhibit, a way to expand upon or highlight certain areas of interest. And a short, engrossing video could be an excellent accompaniment to a grant proposal.

With the use of video technology, one can reach a wide audience beyond an individual institution. An informative and entertaining presentation can capture the attention of researchers, students, administrators, and potential donors. It is always profitable for academic libraries to promote their collections in creative, convincing ways. A well-produced videotape can provide stunning results.

Suggested Reading

Hausman, Carl. 1991. *Institutional Video: Planning, Budgeting, Production, and Evaluation*. Belmont, Calif.: Wadsworth Publishing.

McQuillin, Lon. 1987. *The Video Production Guide*. Ed. Charles Bengsin-

ger. Indianapolis: Howard W. Sams Co.

Schihl, Robert J. 1989. *Single Camera Video: From Concept to Edited Master*. Boston: Focal Press.

Schorzman, Terri A. 1994. "The Library as Producer: The Creation of Local and Personal History Video Collections." In *Video Collection De-*

velopment in Multi-type Libraries, ed. Gary P. Handman, pp. 303-14. Westport, Conn.: Greenwood.

Smith, David L. 1991. *Video Communication*. Belmont, Calif.: Wadsworth Publishing.

Utz, Peter. 1987. *Today's Video: Equipment, Setup, and Production*. Englewood Cliffs, N.J.

Whittaker, Ron. 1989. *Video Field Production*. Mountain View, Calif.: Mayfield Publishing.

Winston, Brian, and Julia Keydel. 1986. *Working with Video: A Comprehensive Guide to the World of Video Production*. New York: AMPHOTO.

Planning for Sustainable Automation

Bruce Harley

The potential for organizational instability should always be considered when proposed library functions and services are reviewed, especially when automation is involved. Sustainable automation can help prevent organizational instability. Planning for sustainable automation requires an environmental perspective, appropriate strategies, and a framework of ecological values.

Organizational instability in libraries can result from organizational change. In the first part of this paper, the relationship between organizational change and instability is explored. Automation has a significant impact on library organization. In the second part of this paper, possible consequences of this impact are contrasted. Planning for sustainable library automation focuses on library organization. In the third part of this paper, the key elements of this planning style—an environmental perspective, appropriate strategies,

and a framework of ecological values—are described.

Organizational Change and Instability

Change in library organization can be an unforeseen result of policies, plans, and procedures. Unforeseen change can be a disruption to library personnel working to achieve shared goals and objectives. The odds for disruption in functions and services increase the more fundamental the change. When this disruption occurs, organizational instability may result.

Library organizational change is not likely to happen simply for the sake of change; instead, change is more likely to be associated with projects that introduce new functions and services. Still, when projects are reviewed, organizational change rarely receives the attention it warrants. This inattention can be intentional. Project reviewers can exclude discussion of organizational change from their deliberations, claiming that it is unlikely or unnecessary. After all, organizational change can require changing an organization's basic assumptions—assumptions that serve to define that organization's culture and mission.¹

Such change cannot be accomplished overnight and may seem too costly and too time-consuming. Nevertheless, failure to consider organizational change can result in serious disruptions to library func-

tions and services during project implementation. Organizational factors to which these disruptions can be attributed include: conflicting goals, functions, and services; outdated, inflexible organizational structure; too few or too many decision-makers; inefficient channels of communication; insufficient allocation of personnel; poorly matched roles and skills; inadequate training programs; inadequate support (clerical, technical, administrative); inefficient workflows; and inaccurate accounting and budgeting.

Such factors need to be identified during the initial consideration given new projects. Projects must be designed to avoid disruptions; otherwise, resulting disruptions to library functions and services can be significant enough to lead to organizational instability. This instability can be more serious when a new project involves automation, given the impact of automation on libraries.

The Effect of Automation on Library Organization

Computers and telecommunications have profound effects on the quantity, distance, and speed of information transfer. In turn, more information moving farther and faster has profound effects on how each of us processes information, perceives the world, and learns. Library personnel, like so many of the users they serve, have experienced significant change in the workplace due to tech-

Bruce Harley is Senior Assistant Librarian, Government Publications & Maps Division, San Diego State University Library, San Diego, California. His e-mail address is harley@mail.sdsu.edu.

nological innovations in information transfer. These innovations have not eliminated all preexisting modes of information transfer. In most libraries, computer workstations still coexist with shelves and tables for books, periodicals, and newspapers. Nevertheless, the rate of introduction and accumulated effect of modern innovations in information technology transcend earlier developments in communications.

A prime example of recent effects of rapid change in information technology is globally networked libraries. A globally networked library is a node on the Internet, linked to libraries around the world. This networked library is especially dependent on computers and telecommunications for access to information. Many critical functions (acquisitions, cataloging, etc.) and services (such as end-user database searching) in the globally networked library are already automated, while other functions and services are candidates for automation.

Given the increasing degree of dependency on automation, can the development of globally networked libraries be described as automation-driven? The answer is yes if those who plan, design, and implement automated library systems fail to concentrate on both the technological and organizational aspects of automation. Furthermore, globally networked library development can be described as automation-driven if library automation planners fail to understand the politics of automation. Automation is capable of changing "patterns of influence and power among key participants in organizations even when no such consequences were . . . intended."² These failures by library automation planners are mutually reinforcing. The more one concentrates on the technological aspects of automation, the more one fails to understand the politics of automation, and vice-versa.

An automation-driven library need not be on the cutting edge of information technology. For a library to be automation-driven, it need only have its own technocracy. In a technocratic workplace the most influential and powerful organizational participants are the personnel with technical expertise. A technocratic workplace is characterized by a rigid organizational structure, centralized decision making, and restricted interdependency and communication among its personnel. Fortunately, sole reliance on technical expertise need not be an inevitable outcome of planning for and implementing library automation projects. Nontechnical expertise can be duly considered, provided library automation planners: give organizational change the attention it warrants—however unlikely, unnecessary, costly, or time-consuming it may seem at the time; identify organizational factors that can cause disruptions to library functions and services; and design projects to avoid disruptions.

A library automation project guided by these principles can lead to library organizational change characterized by broadened interpersonal relationships, shared responsibility in the decision-making process, less hierarchical organizational structure, and redefined authority systems.³ In this context, "redefined authority" systems means a narrowing of the "gulf between library managers and the people they have supervised."⁴ Library change characterized by these features can help ensure a stable organizational structure.

Library automation planners must consider the possible effects of automation on organizational structure. They need to acknowledge the extent to which potentially unstable, inherently fragile automated systems can render their organizational environment equally unstable and fragile.⁵ Library automation plan-

ners must emphasize preventing organizational instability rather than reacting to it. To do so, they must plan for sustainable automation.

Planning for Sustainable Library Automation

Library personnel who plan for sustainable automation are committed to preventing disruptions to library functions and services. The planners' commitment is based on the shared perceptions described below: an environmental perspective, appropriate strategies, and a framework of ecological values. Shared perceptions are vital to a library organization. An organization functions best "when [its] members share certain perceptions: They must perceive that they have common goals . . . and they must have a common perception of the organizational members' roles."⁶ The shared perceptions of those engaged in planning for sustainable automation can serve as the defining assumptions for a library's organizational culture and mission.

An Environmental Perspective

Planning can be environmental. Planning can create "opportunities for staff at all levels of [an] organization to inform themselves about the organization and its environment."⁷ An environmental scanning component of planning has been described as a means of highlighting organizational goals and objectives as well as identifying opportunities to achieve these goals and objectives.⁸ To scan a library organization's environment is to assess the demands and needs confronting the organization, examine the organization's strengths and weaknesses, and address potential problems arising from the assessment and examination.

By scanning their organizational environment, library automation

planners can become aware of what and who comprises this environment. This awareness can provide insights into environmental diversity. In addition, automation planners can gain a greater appreciation for the interrelationships within the library's environment. This appreciation can provide insights into environmental complexity. In the end, the environmental scan can provide library automation planners with a realistic analysis of the physical, social, political, and economic factors that characterize their environment. Planners must rely on this analysis while designing appropriate strategies for library automation projects.

Appropriate Strategies

Planning appropriate strategies that facilitate the appropriate application of technology is essential to successful automation projects. Evaluating the appropriateness of a planning strategy for a library automation project is not simple. Difficult questions must be answered.

1. What is the purpose of the project?
2. Why is the project necessary?
3. Where is the project site?
4. How will the project be implemented?
5. Who will perform the project tasks?
6. Will training and retraining be necessary?
7. Which material resources are needed for the project?
8. When will the project be completed?
9. How much will the project cost?
10. What plans need to be made for project contingencies?

The answers to these questions can help library automation planners realize that "A good strategy is not synonymous with a doable one. Nor is a doable strategy synonymous with a good one. The challenge is to find a good doable strat-

egy."⁹ The answers to these questions can also help convince library automation planners that there is much more to an appropriate application of technology than simply the right hardware and technical skills. An appropriate library automation strategy represents "a frame of reference and a way of thinking."¹⁰ In other words, such a strategy consists of decisions based on shared values.

A Framework of Ecological Values

Values cannot exist in isolation. In order to withstand the tests of time, values require an ideological framework. The concept of sustainability is based on a framework of ecological values. This framework has much to offer library automation planners who are attempting to apply technology appropriately.

In the context of this paper, the concept of sustainability is similar to the idea of sustainable development, popularized during the 1992 United Nations Conference on Environment and Development (UNCED) or "Earth Summit" in Rio de Janeiro. Sustainable development pertains to "development activities that meet current needs without jeopardizing the welfare of future generations by compromising the world's natural ecosystem."¹¹ However, two different viewpoints of sustainability have been espoused: technological sustainability and ecological sustainability. The difference between these viewpoints is "whether a society can become sustainable within the modern paradigm through better technologies and more accurate prices, or whether sustainability requires the transition to a postmodern world that transcends... 'individualism, anthropocentrism, patriarchy.'"¹² Whereas technological sustainability is only hardware- (or technically) oriented, ecological sustainability is both hardware- and software- (or socially) oriented.¹³

Although the environment confronting planners engaged in library automation is neither the entire world nor completely "natural," they must plan ecologically, not just technically. To do so, library automation planners must adopt "five instrumental values... wholeness, posterity, smallness, quality, and community."¹⁴ To value wholeness is to see the big picture and avoid losing sight of the forest for the trees. To value posterity is to create a vision of what the future should be like. To value smallness is to think in terms of appropriate scale. To value quality is to deemphasize quantity for quantity's sake. To value community is to appreciate organizational culture.

Adopted individually, these instrumental values can be put to good use by anyone. However, it is when these instrumental values are adopted by an organization such as a library that the ideological framework of ecological sustainability manifests itself. One example of the manifestation of ecological sustainability is the "bioregional paradigm," which emphasizes region and community vs. state and nation; conservation, stability, self-sufficiency, and cooperation vs. exploitation, world economy, and competition; decentralization, complementarity, and diversity vs. centralization, hierarchy, and uniformity; and symbiosis, evolution, and diversity vs. polarization, uncontrolled growth, and monoculture.¹⁵

Another example of the manifestation of ecological sustainability is the Green Party platform, which emphasizes ecological wisdom, grassroots democracy, personal and social responsibility, nonviolence, decentralization, community-based economics, postpatriarchal values, respect for diversity, global responsibility, and future focus.¹⁶

Neither the "bioregional paradigm" nor the Green Party platform was intended to be applied to library

organizational issues. Nevertheless, some principles common to both value frameworks can be applied to library automation planning. These principles include efficient and effective resource allocation, organizational responsibility and commitment, appropriately scaled technology, cooperation, genuine assistance to all, and future focus.

When adhered to by automation planners, these principles can help foster an ecologically sustainable library organization. Such an organization is not likely to become unstable.

Conclusion

Organizational change in libraries can be a disruptive force, especially when automation is involved. However, the impact of automation on library organization can be beneficial. For this impact to be beneficial, library automation must be sustain-

able. To help ensure that library automation is sustainable, planners must share certain organizational perceptions. These perceptions are an environmental perspective, appropriate strategies, and a framework of ecological values.

References

1. W. Edward Stead and Jean Gardner Stead, *Management for a Small Planet: Strategic Decision Making and the Environment* (Newbury Park, Calif.: Sage Publications, 1992), 187.
2. Peggy Johnson, *Automation and Organizational Change in Libraries* (Boston: G. K. Hall, 1991), 91.
3. *Ibid.*, 136.
4. *Ibid.*, 138.
5. Mary E. Clark, *Ariadne's Thread: The Search for New Modes of Thinking* (New York: St. Martin's Press, 1989), 484.
6. Stead and Stead, *Management for a Small Planet*, 65.
7. Meredith Butler and Hiram Davis,

most commonly among those for reference materials and serials where host library holdings were often missing. Even for records present, call numbers were often missing, multi-volume records were often incomplete, and records for theses and dissertations were confusing and misleading. A dirty database requiring some clean-up of past cataloging sins had been expected. However, the magnitude of the problem and the apparent overlaying of first-loaded records by subsequently-loaded records was totally unexpected by cataloging staff, thereby calling into question the process by which the shared database was created.

In 1984, the University of Rhode Island (URI), Roger Williams University (RWU), Rhode Island College (RIC), and the Community College of Rhode Island (CCRI) formed a consortium called the Higher Educa-

"Strategic Planning as a Catalyst for Change in the 1990's," *College & Research Libraries* 53 (Sept. 1992): 397.

8. *Ibid.*, 398.

9. Robert H. Waterman, Jr., "The Seven Elements of Strategic Fit," *The Journal of Business Strategy* 2 (Winter 1982): 69.

10. Association of Research Libraries, *Strategic Planning in ARL Libraries*, SPEC Flyer #108 (Washington, D.C.: Association of Research Libraries, 1984).

11. Julie White, "Thinking Small: U.N. Looks to Islands for 'Sustainable' Model," *The Interdependent* 20 (Summer 1994): 1.

12. David W. Orr, *Ecological Literacy: Education and the Transition to a Postmodern World* (Albany, N.Y.: State University of New York Press, 1992), 24.

13. *Ibid.*, 181.

14. Stead and Stead, *Management for a Small Planet*, 129.

15. Kirkpatrick Sale, *Dwellers in the Land: The Bioregional Vision* (San Francisco: Sierra Club Books, 1985).

16. Charlene Spretnak and Fritjof Capra, *Green Politics* (Santa Fe: Bear & Company), 230-33.

tion Library Information Network (HELIN). Providence College (PC) subsequently joined the consortium in March of 1992. The goals of the consortium were to make available the resources of the state's academic libraries to the students of the participating institutions and the citizens of the state, share the costs of an online catalog, and save taxpayers money since three of the original four libraries were public institutions.

Each library in the consortium was given equal voting status in the policymaking process to prevent the smaller libraries from being overpowered by URI, which was not only the largest library, but the only library with a systems librarian and the host site for the future consortium database. HELIN committees were formed for cataloging, reference, circulation, and reserves to recommend policy for the consortium. Final decisions on consortium mat-

OPAC Database Creation Problems

Nadine L. Baer, James A. Barrett, and Karl E. Johnson

This article reports on a study to determine the nature and extent of the problems discovered when cataloging records for a consortium of five libraries were merged to create an online public access catalog (OPAC). Records in the host (University of Rhode Island) library's shelflist were compared to records in both the Online Catalog Library Center (OCLC) database and the OPAC. Problems were most acute in shared records,

Nadine L. Baer is Serials Cataloger, James A. Barrett is Monographic Cataloger, and Karl E. Johnson is Monographic Cataloger at the University of Rhode Island.

ters were and are still made by the management committee, composed of the directors of the member libraries. A HELIN committee was formed to select an automated system for the consortium and, after much deliberation, Innovative Interfaces, Inc. (III) was chosen. A grant was obtained from the Champlin Foundation to offset the cost of the central processing unit.

In March of 1991, the staff of the URI Library received access to its new online catalog, nine months before making it available to students. Along with this access came a rumor that 20,000 main entry records beginning with the letters R, S, and T had somehow been lost during the process of record conversion and were missing from the database. It soon became obvious to inquisitive catalogers that strange mutations, often weird and not so wonderful, had also occurred to some data. One oddity soon discovered was the existence of duplicate data: occasionally as many as eight to ten records for one item. Then came the discovery of missing call numbers and missing holdings, often made more confusing by the fact that bibliographic records attached to item records of affiliated institutions sometimes indicated that URI, whose records had been loaded first, was the originator of the bibliographic records. URI AACR2 serial records were often found lacking, but non-AACR2 serial records from another consortium member were present. Next, the plight of cataloging records for theses and dissertations came to light. Here it was not uncommon to have only one of two item records (stacks and archives) for a master's thesis or one of three item records (stacks, archives, and microfilm) for a doctoral dissertation to be present in the OPAC database. In addition to these problems came evidence that the OPAC had been misindexed, resulting in unused as well as authoritative series entries being present in the

OPAC along with their respective "See" references. Clearly, something was wrong!

Catalogers soon were called to resolve inconsistencies in records that were brought to light via the circulation process. Circulation staff often could not correlate books brought to the circulation desk with their corresponding bibliographic records, in part because the bibliographic records were missing and in part because many items lacked barcodes. In response, additional records were created containing only call number, author, and title. The cleanup process was slow and tedious—and appeared overwhelming.

It was with this background that the authors, having become curious as to the nature of the new catalog, decided to investigate the magnitude of the problem. The decision was made to compare the OPAC database with URI's main shelflist and the OCLC database, as the cataloging records used to create the OPAC database had been created electronically via OCLC since 1972 and older records had been converted retrospectively by OCLC beginning in 1987. These electronically created records had been tape-loaded into a C L Systems, Inc. (CLSI) circulation system to create a database where item records for library materials were created at the time of circulation. The resulting database was extracted by Autographics, corrected for subject heading changes by Blackwell, and finally loaded into the III OPAC database. Subsequently, records for other consortium members were loaded into the database. Before loading records for CCRI, RIC, and RWU, a decision to allow subsequently loaded bibliographic records accompanied by item records to overlay previously loaded bibliographic records lacking item records was made. This decision was reversed before loading records for PC, after recognition that much damage had been done to the database.

Literature Search

An online search was conducted in *Library Literature* for previous investigations regarding this topic. Except for five studies concerned with errors within electronic bibliographic records,¹ nothing was found.

Methodology

The authors decided to compare records in the URI shelflist with the OCLC database and the OPAC. Sampling was conducted by using shelflist drawers chosen from Library of Congress (LC) subject classifications in the collection to facilitate record keeping. While a more scientific method could have been used, this method was determined to be adequate.

The main shelflist at URI contains approximately 510,000 monographic records distributed among 611 drawers. One drawer from each of the larger URI LC subject classifications was chosen for examination. A shelflist inventory form (see figure 1) was used to record call number, date of publication, OCLC control number, and other relevant data for each record examined. The number of records missing from each database and the number of OPAC records missing call numbers or barcodes, or shared with other consortium members, were tallied. Other problems were noted and subsequently tallied if found significant.

A total of seventeen drawers containing 14,174 records were examined and compared first to the OCLC database and then to the OPAC database (see table 1). Simple corrections, such as the insertion of holding libraries and call numbers, were made to the OPAC records during examination, but little effort was made to compare physical items with records or to retrieve items for barcode attachment. In general the shelflist was assumed to be correct,

Table 1
Record Summary

Drawer #	Classification	Missing/OCLC	Missing OPAC	MSG Call #	MSG Barcode	SR %	MSG/SR
9	B3501-BC455.4	10/952=1%	70/952=7.4%	37/882=4.2%	206/913=23%	419/952=44%	61/419=15%
36	D550-D764.6	16/909=1.8%	94/909=10.3%	155/815=19%	323/1019=32%	493/909=54%	35/493=7.1%
71	E185.8-E189	4/802=0.5%	113/802=14.1%	45/689=6.5%	149/798=19%	528/802=66%	91/528=17%
86	F1035-F1234	13/919=1.4%	112/919=12.2%	13/906=1.4%	320/919=33%	374/919=41%	90/374=24%
95	GC127-GF895	30/921=3.3%	159/921=17.3%	49/762=6.4%	160/810=21%	523/921=57%	143/523=27%
160	HE8870-HF1025	9/796=1.1%	79/796=9.9%	50/717=7%	241/749=32%	275/796=35%	58/275=21%
222	JK2541-JN199	20/877=2.3%	110/877=12.5%	123/757=1.6%	359/828=43%	413/877=47%	87/413=21%
237	KF1135-KF2995	17/649=2%	78/649=12%	18/632=2.9%	106/611=17%	253/649=39%	51/253=20%
255	LB2325-LB2342	13/872=1.5%	122/872=14%	33/750=4.4%	112/784=14%	508/872=58%	103/508=20%
273	ML420-ML1993	14/909=1.5%	83/909=9.1%	143/826=17%	215/915=23%	427/909=47%	75/427=18%
288	NC15-NC7572	20/896=2.2%	87/896=10.8%	33/799=4.1%	194/863=23%	337/896=38%	67/337=20%
393	PR6068-PR8727	9/776=1.2%	75/776=9.7%	49/701=7%	241/802=30%	248/776=32%	36/248=15%
470	QD320-QD434	17/776=2.2%	48/766=6.3%	36/718=5%	258/884=29%	78/766=10%	10/78=12.8%
524	RC480.5-RC499	18/919=2%	88/919=9.6%	51/831=6.1%	103/885=12%	432/919=47%	76/432=18%
543	SB470-SB945	9/687=1.3%	46/687=6.7%	35/641=5.5%	142/690=21%	174/687=25%	29/174=17%
564	TJ214-TJ1078	11/837=1.3%	33/837=3.9%	70/804=8.7%	242/850=29%	114/837=14%	19/114=17%
599	Z1025-Z1209	15/687=2.2%	97/687=14.1%	127/590=22%	485/988=49%	264/687=38%	75/264=28%
(17)	Average %	1.7%	10.6%	8.4%	26.5%	40.7%	17.9%

missing records (13.6%) within the letter S, but missing R and T records totaled only 7.9% and 4.6%, respectively. In addition, the letters A through C, and M, V, and Y showed missing records ranging from 6.9% to 7.8%, with other letters following with lesser values (see table 2). These numbers suggest that the long-ruled R, S, and T missing records consist essentially of main entry records beginning with the letter S.

Approximately 8.4% of the records in the OPAC were lacking call numbers. Results ranged from a low of 1.4% in drawer #86 (History—Canada/History—Mexico) to a high of 22% in drawer #599 (General Bib-

liography/National Bibliography—America). The overall percentage of records lacking call numbers (8.4%) is low if one considers accuracy of holding library designations as part of the call number. Examination of OVERSIZE designations in drawer #288 (Drawing - Design - Illustration) shows the OVERSIZE prefix missing from 57 out of 228 records (25%) and a companion study conducted within the library's reference collection shows the REF. designation missing from 559 out of 3,243 records (17.2%).³

Approximately 26.5% of the URI item records lacked barcodes. Results ranged from a low of 12% in

drawer #524 (Neurology and Psychiatry) to a high of 49% in drawer #599 (General Bibliography/National Bibliography—America). This is not surprising considering there was no concerted effort to barcode the entire collection before database loading. In fact, considering that many items were barcoded at time of circulation, this figure may be an indication of relatively high collection circulation.

Approximately 65% of the total thesis/dissertation records were incomplete, apparently due to inadequate input procedures. As mentioned previously, it was common to find only one of two records (stacks and archives) for a master's thesis or

Table 2
Missing Alphabetical Entries

Drawer	9	36	71	86	95	160	222	237	255	273	288	393	470	524	543	564	599	Total	(%)
A	2	8	11	8	8	1	15	9	7	13	8	10	4	5	2	2	6	119	(7.7)
B	4	14	7	9	16	8	6	2	10	3	10	0	4	5	1	5	10	114	(7.2)
C	6	10	4	6	11	4	18	5	10	5	5	7	5	1	4	0	5	106	(6.9)
D	2	4	2	5	6	2	1	2	5	2	1	2	2	4	0	0	4	44	(2.9)
E	3	1	3	2	7	3	4	4	7	0	1	3	1	3	3	0	5	50	(3.3)
F	5	5	7	3	5	3	5	4	5	3	2	1	3	5	1	0	1	58	(3.8)
G	1	17	8	5	4	3	7	4	4	3	5	2	1	7	2	0	3	76	(4.9)
H	1	3	8	4	8	3	3	4	4	3	9	9	0	4	2	2	5	72	(4.7)
I	1	4	0	3	0	2	1	0	1	0	0	9	1	1	1	0	1	25	(1.6)
J	1	2	1	4	2	2	0	2	2	1	1	1	0	4	0	0	0	23	(1.5)
K	0	0	4	5	5	2	2	2	7	2	1	0	0	2	2	0	3	37	(2.4)
L	6	10	7	3	9	2	1	2	5	1	1	0	1	2	0	1	2	53	(3.6)
M	13	11	5	7	13	8	8	3	10	5	6	2	3	2	4	1	10	111	(7.2)
N	3	4	4	2	1	0	4	3	3	1	1	1	1	3	1	1	6	39	(2.5)
O	1	0	1	2	6	1	0	1	2	0	2	6	0	2	1	0	1	26	(1.7)
P	4	5	3	2	9	1	5	0	5	1	7	0	1	5	3	0	5	56	(3.5)
Q	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	2	(1.3)
R	8	5	10	10	10	9	7	9	6	11	16	5	2	6	6	0	2	122	(7.9)
S	13	18	8	20	22	14	10	17	9	20	13	2	13	22	6	2	0	209	(13.6)
T	4	9	4	7	7	6	3	1	4	4	2	7	1	1	1	1	8	70	(4.6)
U	0	2	3	1	0	0	0	0	1	0	0	15	0	0	2	0	2	26	(1.7)
V	0	0	3	2	2	0	0	1	0	1	0	0	0	0	0	1	1	11	(7.2)
W	2	4	7	2	10	4	3	2	4	4	5	0	3	3	4	5	6	68	(4.4)
X	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	(0)
Y	0	3	0	0	2	0	1	0	0	0	0	4	0	1	0	0	1	12	(7.8)
Z	0	0	0	1	0	1	1	0	1	0	0	1	2	0	0	0	0	7	(4.6)
Total	80	139	110	113	161	79	105	78	113	83	97	87	48	88	46	21	88	1,536	(100)

only one of three records (stacks, archives, and microfilm) for a doctoral dissertation. In addition, the institutional/department-specific author/series statement (which, incidentally, was to be found on the verso of the input form) was frequently lacking.

Approximately 65% of the multi-copy/volume records were also incorrect. This was to be expected, as record updating had been done manually for some years—only in the public catalog—in order to save the cost of electronic updating. In addition, the URI library's holdings were missing from approximately 17.9% of the 40.7% of the bibliographic records shared with other consortium libraries. There is a relationship between the shared record percentage and the percentage of records missing from the OPAC, as most drawers with a high percentage of missing records also have a high percentage of shared records. Drawers #71, #95, #255, and #599, with MISSING/OPAC records of 14.1%, 17.3%, 14%, and 14.1%, respectively, have shared record percentages (SR. %) of 66%, 57%, 58%, and 38%, respectively. In addition, drawers #36, #86, #222, #237, #288, #393, and #524—with MISSING/OPAC records of 10.3%, 12.2%, 12.5%, 12%, 10.8%, 9.7%, and 9.6%—have shared record percentages of 54%, 41%, 47%, 39%, 38%, 32%, and 47%, respectively. Conversely, drawers #470 and #564, with MISSING/OPAC records of 6.3% and 3.9%, have shared record percentages of only 10% and 14%, respectively.

The companion study conducted within the reference collection and referred to above found 4,326 records missing out of a total of 19,303 (22.4%). Of the 4,326 records found missing, 2,295 (53%) were shared with other consortium members. This high percentage of missing reference items from shared records is not surprising when one considers that, because reference items do not circulate, item records could not be

made at time of circulation. Invariably, when URI holdings were present on shared records, either a call number or a barcode or both were present on the URI item record, leading to the inescapable conclusion that, for those missing or overlaid URI holdings, even minimal item records had never been created.

Conclusions

Merging the four original databases, with their individual idiosyncrasies and without comprehension of possible consequences, led to significant corruption of the URI holdings in the OPAC database. Analysis of the data suggests that item records (minimal at least) should have been created for all URI bibliographic records before loading other consortium member library records into the online catalog. This would have prevented the overlay of first-loaded URI bibliographic records that lacked item records by subsequently loaded consortium member records accompanied by item records. The decision to allow overlay also resulted in the replacement of a number of URI AACR2 serial records containing specific holding notes by non-AACR2 records. Much trouble could have been avoided if all items in the collection had been barcoded and related to their bibliographic records before loading the records into the online catalog. Non-barcoded items forced the creation of temporary records, thereby degrading the quality of the database.

The presence of item records containing barcodes but lacking call numbers was apparently the result of the shortsighted construction of a database established solely for circulation purposes without consideration to the future needs of an OPAC. Initial ignorance regarding the need to distinguish between holdings in main and subsidiary libraries, as well as specific collections within li-

braries, has led to the need to manually adjust location fields for over 90% of the URI records in the database. More careful attention to procedures and training when inputting records could have prevented the myriad of problems found with the thesis and dissertation records. The unwise decision made long ago to save money by updating multi-volume records manually in the public catalog rather than updating them electronically via the OCLC database apparently led to the many discrepancies in those records.

If there is a lesson to be found in this analysis, it is that foresight in the planning process, as well as careful attention to procedures, are essential in any endeavor—but are especially so when producing records to be used in an electronic catalog. In a consortium setting with a shared database, the process is even more treacherous and greater care must be taken unless one wishes to be bitten in the backside by the sharp teeth of ignorance and inexperience.

References

1. James Gerald Schoenung, "The Quality of Member-Input Monographic Records in the OCLC On-Line Union Catalog" (Ph.D. dissertation, Drexel University, 1981); Jaye Bausser, Jimmie Y. Davis, and David Gleim, "Error Detection in Bibliographic Records: Can the Computer Do It All?" in *Energies for Transition* (Chicago: Association of College and Research Libraries, 1986), 27-29; Mary Ellen Soper, "What You See May Not Be What You Get: Errors in Online Bibliographic Records for Serials (in the OCLC Database)" *The Reference Librarian* 27-28 (1989): 185-213; Peter Jacso, "Searching for Skeletons in the Database Cupboard; Part I: Errors of Omission" *Database* 16, no. 1 (February 1993): 38-49; and Pamela Cahn, "Testing Database Quality" *Database* 17, no. 1 (February 1994): 23-30.
2. Mimi Keefe, reference bibliographer, University of Rhode Island. Unpublished data, 1994.
3. Ibid.

Migration from LS/2000 to KeyNOTIS: A Medium-Sized Academic Library's Experience

Joyce L. Huang
and Ting Zheng

The University of Wisconsin-Whitewater Library migrated from one automated system (LS/2000) to a second-generation integrated automation system (KeyNOTIS). This article describes the terminal network setup and public terminals setup for the new system utilizing existing equipment. It also describes database conversion problems that were encountered and how clean up work proceeded. The authors provide information and advice for those about to embark on a similar migration project.

Since Ameritech's announcement in 1991 that it will no longer support the LS/2000 system after July 1995, LS/2000 users have faced the inevitability of migrating to a new automated system. The University of Wisconsin (UW) at Whitewater was designated by the UW System as one of the four sites to lead the migration from the LS/2000 to the KeyNOTIS automated system, a turnkey version of NOTIS. Consequently, the UW-Whitewater Library administration created two teams to deal with the migration issues. An automation implementation team was

Joyce L. Huang is Head of Technical Processing Service at the University of Wisconsin-Whitewater Library. Ting Zheng formerly was Coordinator of Information Technology at the University of Wisconsin-Whitewater Library. Currently he is with TDS TELECOM, Madison, Wisconsin.

selected from the library and the computer center to address issues of hardware and system implementation, and an automation committee from the library was asked to address profiling, preparing for the conversion, testing the converted database, and cleaning up the converted database.

LS/2000 was an OCLC local system that was sold to Ameritech. As a turnkey multifunction integrated automated library system, it runs on Data General Eclipse and Eagle, DEC PDP/11 and VAX, and IBM Series I hardware. The system supports local cataloging, authority control, circulation control, a patron access catalog, and interlibrary loan. Serials control and acquisitions functions are supported on standalone PC-based systems—the SC350 serials control system and the ACQ350 acquisition system. The link to LS/2000 allows the workstation full access to all LS/2000 functionality, but does not allow one to perform 350 functions from LS/2000.

KeyNOTIS is a turnkey version of the NOTIS library automation software package developed by Northwestern University Library and later sold to Ameritech. It utilizes IBM hardware and operating systems and provides integrated support for public access catalog, acquisitions, serials control and claiming, circulation control, cataloging, authority record management, and data base management. The strength of the system is in its Boolean searching capabilities.

A Beta test of KeyNOTIS was conducted in 1990 at California State University-San Bernardino.¹ Since then, 17 KeyNOTIS sites have been established in the United States, including the four University of Wisconsin sites. A search of current literature shows few articles on migration between systems.² Only the Ameritech Information Systems' *Lines of Support* has published some short notes on migration issues.³

This article describes one migration experience; other libraries planning similar moves may find the information and advice valuable. This article focuses on three areas: network setup selection, conversion, and cleanup work.

Terminal Network Setup Selection

The KeyNOTIS system installed at UW-Whitewater operates on VSE/ESA software on a shared IBM ES9121 computer. The design of our KeyNOTIS network was dictated somewhat by a number of factors, including the following:

- The full utilization of an existing ADM11 program for circulation and OPAC functions and Windows-based PC workstations for access to other library resources in the state and nationally.
- The use of existing LS/2000 cables (RS232 cables). The Whitewater campus is in the process of installing a campuswide fiber optic backbone that will allow us to connect to the campus Ethernet network without cables.
- To minimize the usage of the campus Gandolf data switch network that only supports a maximum of 9600 bps.

In planning for terminal connection to an IBM mainframe, we had two options. The first was a 7171 connection, an IBM ASCII terminal server that supports ADM, VT100, and other terminal emulations. This is the traditional IBM terminal connection. The advantage of this connection is that it provides each terminal a CICS (Customer Information Control System) address that gives NOTIS easy control of its terminals. However, the 7171 supports only one session per port and limits a terminal's direct access to a TCP/IP network. This option locks

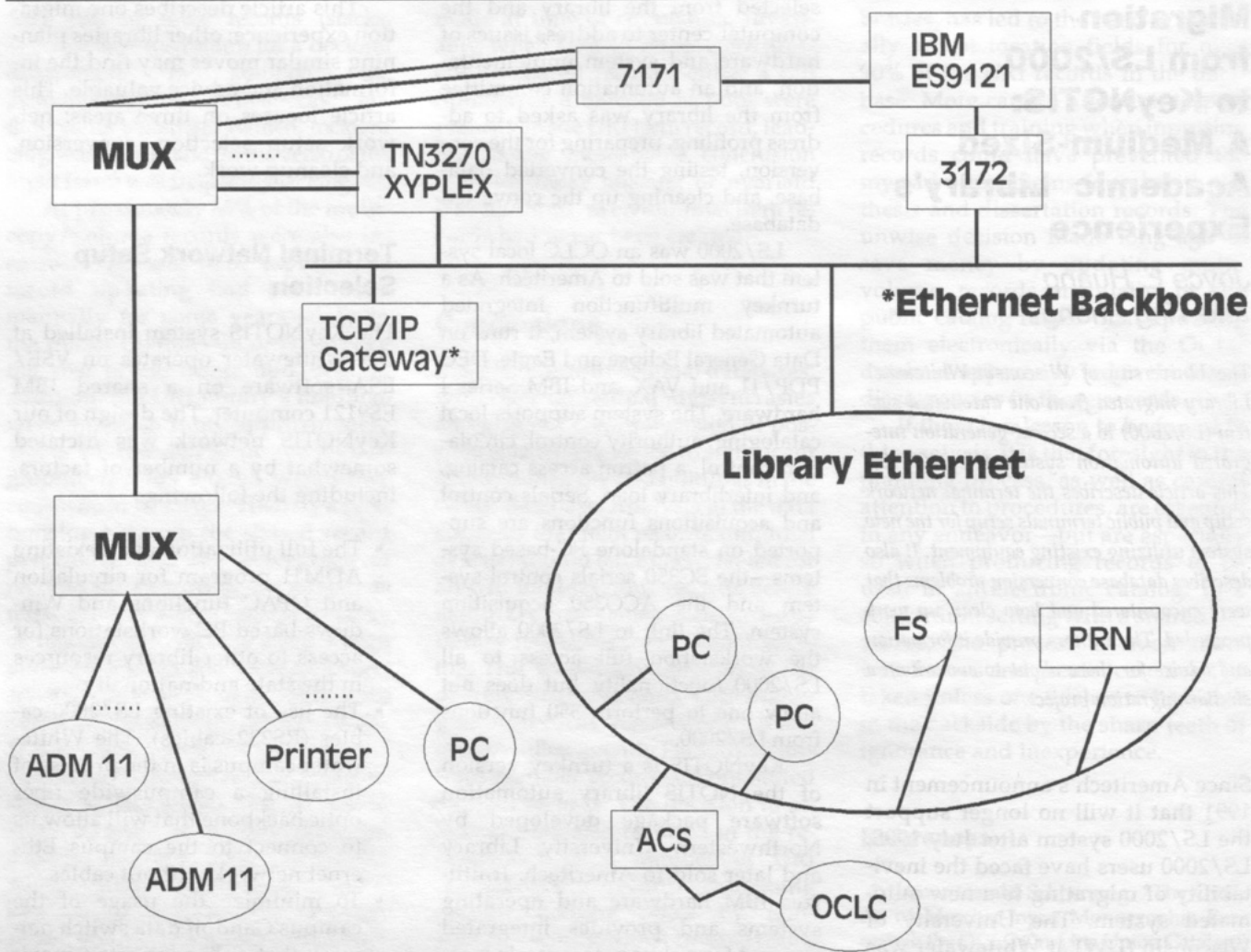


Figure 1
UWW KeyNOTIS Network

ADM11 = ADM11 dumb terminal
 PC = personal computer
 ACS = asynchronous communications server
 7171 = IBM asynchronous terminal server

FS = file server
 PRN = printer
 XYPLEX = XYPLEX terminal server with TN3270 support

* = To be implemented. The library PCs now are connected to IBM via Gandalf master switch.

us into IBM technology and incurs a high maintenance cost.

The second option was using XYPLEX, 3COM, or DATABILITY terminal servers, which allows TCP/IP communication via the TN3270 protocol. Advantages of this connection include support of multiple simultaneous sessions per port and

easy access to a TCP/IP network. A disadvantage is that an individual CICS address is not possible for each terminal. We chose the XYPLEX terminal server because it supports TN3270 and multiple simultaneous sessions per port, is compatible with the existing LS/2000 cables (RS232), supports standard key mapping for

both ADM and VT100 emulation, and, with one connection, provides the user access to IBM, VAX, and other servers.

Figure 1 illustrates the current KeyNOTIS network connection at UW-Whitewater.

The performance of this network partially depends on the efficacy of

the existing multiplexers, some of which support speeds only up to 4800 bauds. They support "transmitter on/transmitter off" (XON/XOFF) but not "data terminal ready" (DTR), which means that terminals have to be reinitiated after a "power off."

OPAC and most staff terminals are connected to the network via XYPLEX. The staff module of KeyNOTIS is accessible only by password. Certain NOTIS functions, such as circulation check-in/check-out and printing, require a terminal CICS address predefined in the NOTIS terminal table (NSYS). As a temporary solution, circulation terminals are connected to an IBM ES9121 via an IBM 7171 terminal server so that each terminal has a CICS address and can be defined in NSYS. Meanwhile, XYPLEX terminals are grouped so that they can be defined in NSYS, a NOTIS transaction for system control (e.g., 1\$\$\$ to define CICS address starting from one).

MDAS and Public Terminals

NOTIS MDAS allows commercial databases to be mounted on the system computer and searched in the NOTIS environment. MDAS allows database searches to be downloaded or printed on the printer associated with the terminal. Because of wiring limitations and difficulty in establishing CICS addresses for terminals, the MDAS print capability was turned off. Patrons using PCs can print the screen or download to a file.

PCs for the public area have to be secured by disabling the control break key using a shareware program called BRK.COM and BOOT-LOCK.COM.⁴ Pibterm, a freeware communications software used for NOTIS access via a XYPLEX terminal server, is programmed to automatically connect to the NOTIS screen with a time-out feature when a user leaves the terminal. This also

prevents users from exiting to the DOS program.⁵ The Windows PROGMAN.INI file is modified to disallow users from altering Windows settings and to limit users to only two program groups. All library applications are in the LIBRARY.GRP group.⁶

Conversion and Testing of Conversion

Conversion from tapeout to acceptance of the database took about six months. The tapeout program, using the %GT utility from LS/2000, involved 428,093 bibliographic records. It was started in March 1993 and completed in April, during which time the library minimized most of its activities on LS/2000. NOTIS QuickLoad+ was used to convert the LS/2000 bib/holdings/item data. Meanwhile the library automation committee was asked to fill out a questionnaire developed by NOTIS Conversion Services to describe both the library's archival cataloging data and the way in which the library would like that data mapped into the NOTIS record structure. It was important to document and analyze past cataloging practices and transfer them accurately to the questionnaire because NOTIS Conversion Services used that information to formulate the work letter and the customized conversion. The conversion process also allowed us to improve cataloging practices. For example, we requested that QuickLoad+ to do the following in bibliographic processing:

- bracket subfield "h" of the 245 field if it is not already bracketed so that [videorecording] will stand out in title;
- strip subfield "w" from fields: 1xx, 240, 4xx, 6xx, 7xx and 8xx;
- strip subfield "c" from 020 field; and

- check second indicator in certain fields for nonnumerics and verify that they are "0" and, if necessary, correct filing indicators for English-language articles.

All location codes and variations together with material type should be included in bibliographic processing. Locations that were left out of the chart were dumped into a default file and made cleanup difficult. When locations have a mixture of classification types, one should select "other class scheme" rather than LC or Dewey; otherwise, indexing errors will occur for those entries classified differently from the chosen scheme.

Audit report printouts from the conversion should be used for manual cleanup. Records that were rejected as invalid were written to one of two error files: MARC error (ME) records and nonstandard error (SE) records. When LS/2000 records could not be converted to the MARC format, ME was used; when NOTIS could not convert the MARC format to NOTIS file format, SE was used.

The test file we prepared to test the converted database included multiple volume sets; multiple copies; a variety of formats and locations; items added at different periods; and items with parts, numbers, and supplements at the end of the call number. Data for comparison between NOTIS and LS/2000 records included author, title, location, call number, enumeration/chronology (vol./no.), item id (bar code number), charges (checkout and renewal on LS/2000), and browses (internal use).

There were three tape loads for testing, including the one that was finally accepted. The first tape load of the converted database was delivered in early July 1993. Many errors noted were due to location table errors or class code errors that we provided to NOTIS. Some errors were due to NOTIS programming. Item records not transferred to NOTIS re-

sulted in no call numbers in the copy holdings record, location mix-ups, and so on. There was corruption of other data, as well.

The second tape load was mounted in early August. While this database reflected sharp improvement over the previous version, certain major errors remained. These included circulation charges and browse statistics being omitted; corrupted data not being corrected; and a fair number of item records not being transferred.

The third version of tape load was delivered in September. Additional improvements had been made. Problems created from the library's processing inconsistency could not be corrected by a conversion program. Other glitches would require much more time for NOTIS to resolve. Meanwhile, with the beginning of a new academic and a new fiscal year, we wanted entering students to have to learn only one system, and we wanted the library acquisitions and circulation functions to be in the same system. We decided to accept the imperfect database as it was and do the cleanup manually.

Problems from Data Migration and Solutions

NOTIS's treatment of multivolume information created certain problems. First, each LS/2000 item record generated a MHLID (MARC holding) line with volume information in NOTIS. However, the sort was not strictly numeric (e.g., 1,11,2,20, etc.). Consequently, volume information skipped around. Second, when there were multiple copies of a volume, a second copy was created for that volume. Subsequent volumes, although single copies, were carried over to the second copy location. Third, because of the way that data in the item record were mapped into NOTIS, nonstan-

dard coding and inconsistency in data entry in some item records generated split copy information. For example, if a "guide" in a multivolume work was entered in the COPY field in LS/2000, it would be identified as a different copy during the conversion.

During the cleanup process, we used the conversion audit report that listed all records for which a MHLID had been created. Summary holdings statements were created to replace the poorly sorted detail holdings statements. Split copy errors were corrected. More than 16,000 records were involved—3.7% of the collection.

NOTIS conversion allows a maximum of 55 items that can be associated with a particular bibliographic record. Bibliographic records with more than 55 items attached were dropped during the conversion and were coded MARC errors (ME) on error reports. Catalogers had to re-export the records from OCLC, and item information had to be re-created.

Non-MARC bibliographic data caused problems. Many short bibliographic records created on ACQ350 and linked to LS/2000 and records for reserve materials got lost. Outstanding order records were cleaned up and most reserve materials were re-entered.

For multivolume works with different call numbers but housed in the same location, only the first call number would show on public display because only one copy record had been created. Normally this is not a problem, but OECD publications have distinct variant call numbers for different volumes. We are still seeking a solution to this problem.

Item messages such as "lost," "missing," and "housed in..." were not converted into NOTIS. Such messages were printed from LS/2000 reports to be re-entered into NOTIS. Only circulation and browse counts were converted.

The default file consisted of records with: no match in the location table; no item attached to the record or item with invalid length ID, i.e., system-generated ID, (all acquisition records and federal document records that were tape loaded fell into this category); and other records which for some reason did not get the item records mapped. We were able to do "set cat" on KeyNOTIS and print the default file records. We then printed the same records from LS/2000 before we discontinued LS/2000. This facilitated the cleanup. Those records identified to be with bona fide items (about 1,100 records, or about .0002% of the database) were quickly re-created on NOTIS. This resolved the problem of having items with bar codes but no item record for use in circulating materials.

Cataloging on NOTIS began six months before we converted the database, so many acquisition records became duplicate records. These items were received and cataloged. We used ACQ350 print-outs to clean up the duplicate records and correctly annotate the orders that were still outstanding. There were about 3,500 such records.

The document records (about 100,000 records) were too massive to be corrected manually. We decided to order smart bar codes from Marcive and wait for the new GTO version from NOTIS (which will be able to create items via GTO) to reload the document records.

SC350 detailed holdings information was not converted. Instead, working from an SC350 printout of summary holdings, all periodicals summary holdings were keyed in within three weeks of bringing up the system.

Items linked to the wrong bibliographic records were the hardest mistakes to identify. They did not show up on the conversion audit reports but were accidentally discovered when we tried to re-create the bona fide items in the records

"dumped" in the default location. We found only a handful of them. In each case, the item record had been attached to the wrong bibliographic record that did not have an item. Relinking resolved the problem.

With some records, the conversion created two item records with one in the default location. The conversion gave both items the same ID, although the NOTIS system does not accept the same ID on different items. This happened in only one location—the Arts Media Center. Since the extra items were mapped to the default location, they were easily identified and deleted.

The conversion program did not strip the initial article when present on MARC tag 700, subfield t. Because we had failed to identify this problem in the work letter, we decided to correct the 1,100 records manually and avoid extra charges from NOTIS for customized work.

Conclusions

No migration from one system to another is simple. File structures are different among systems. NOTIS, for example, has distinct bibliographic records, copy records, and item records; while, in LS/2000, copy and item records are the same. Acquisition and serials are integrated into one NOTIS system with records attached to the same bibliographic records, but in LS/2000 they are subsystems in separate databases linked to LS/2000 only through 350link.

With a collection of 428,000 bibliographic records and many more item records, some losses during

transfer are bound to happen. In general we were pleased with the conversion. Except for the 16,000 records with MARC holdings information that sorted strangely, corrections were manageable. The transition from the old system to the new was smooth; there was no major disruption of service. Problems that affected service were quickly corrected—for example, providing summary holdings information on periodicals, recreating item records for those inadvertently dumped into a default file, and deleting duplicated acquisition records.

In summary, from the UW-Whitewater experience, a successful migration would require the following five points:

- careful planning of the migration process in a timely manner;
- team work among staff in all service areas who know the needs of each area, and with campus computing experts who provide computing assistance;
- implementing functions such as cataloging several months before conversion helped us identify potential conversion problems and familiarize ourselves with the new system;
- designing a state-of-the-art library network system under existing computing environment to facilitate staff functions and public access; and
- maintaining detailed logs of profiling changes, conversion problems, and tested records so that subsequent testing and corrections can be done efficiently.

References

1. Beverly A. Ryan, "Beta Testing an Integrated Library Automation System," *Information Technology & Libraries* 11 (December 1992): 367-71.
2. Julie Hallmark and C. Rebecca Garcia, "System Migration: Experiences from the Field," *Information Technology and Libraries* 11 (December 1992): 345-58; and Lester J. Pourciau, "Automated Library System Migration in the United States," *Electronic Library* 10, no. 2 (April 1992):103-8.
3. Tari Keller, "Notes on Migration from LS/2000 to NOTIS or Another System," *Lines of Support* (Fall 1992): 3-5.
4. BRK.COM and BOOTLOCK.COM shareware are available from Foley Hi-Tech Systems (172 Amber Drive, San Francisco, CA 94131; (415) 826-6084).
5. Batch file to prevent users from exiting to DOS from Pibterm and to be automatically connected to NOTIS screen:
Set compsec = c:here.not (here.not is a dummy text file)
pibterm/s = notis.scr (start Pibterm with a script file)
set compsec = c:\command.com (reset to command.com after exit from pibterm).
6. Changes we made to the PROGRAM.INI FILE:
[Restrictions]
NoRun = 1
NoClose = 1
NoSaveSettings = 1
NoFileMenu = 1
EdiLevel = 4
[Settings]
Windows = 68 48 580 384 1
Display .drv = vga.drv
Order = 2 1
[Groups]
Group1 = c:\windows\library.grp
Group2 = c:\windows\control.grp

Internet Use through the University of Toronto Library: Demographics, Destinations, and Users' Reactions

Joy Tillotson, Joan Cherry, and Marshall Clinton

Using an online survey, focus groups, and logs of telnet connections, the authors studied users who gain access to the Internet via the University of Toronto Library online catalog and information system. The study showed that 75% of the users were male, 46% of the users reported finding nothing that they were looking for, and 13% of the Internet sites available accounted for 80% of the telnet connections. The paper concludes with a discussion of the implications for professional practice and future research.

Although the Internet is known to have millions of users, very few studies have been published about what it is used for and by whom. The few studies that have been published focus on use by librarians. Tillman and Ladner¹ studied the uses of the Internet by special librarians. McClure, Moen, and Ryan^{2,3} conducted a survey in which they asked librarians in academic and

public libraries to assess the impact of the Internet on libraries. Eisenberg and Milbury⁴ surveyed members of LM-NET: The School Library Media Network, a listserv whose target audience is school library media specialists, about their use of Internet resources and the impact of their use of Internet on the role or status of their library media programs. In addition to formally published studies, some informal information about use of campuswide information systems is available, such as the popularity rankings posted by InfoSlug at the University of California Santa Cruz⁵ and the gopher at Texas A&M.⁶ What is missing are formal studies of Internet users who are not librarians. Since the whole user group of the Internet would be difficult to study, we studied the users at one site.

We looked at people who used the Internet through UTLINK, the online catalog and information system at the University of Toronto. The University of Toronto is a major research university in Ontario, Canada, with an enrollment of fifty thousand students. At the time of the study, students were not automatically eligible for computer accounts that gave access to the Internet, so many of them took advantage of the access to the Internet offered by the library. The research questions we investigated in the study included:

1. Who is using the Internet? Are there differences in sex or academic status between Internet users and the overall population at the university?
2. Where do they go on the Internet? What sites do they access and why?
3. What are they using the Internet for?
4. Are they satisfied with what they find on the Internet?
5. Are they satisfied with the way the University of Toronto provides access to the Internet?

6. What changes would they like to see?

We hoped that the answers to these questions would prove useful to librarians who are planning services related to the Internet.

Methodology

We used an online survey, focus groups, and analysis of a log of telnet connections to address the research questions. The online survey results provided demographic information, information about people's reasons for using the Internet, and users' ratings of the information they found, as well as of the method of providing access. The open questions on the survey and the focus groups were helpful in collecting suggestions for improvement and explaining the uses people made of the Internet. Examination of the logs of telnet connections helped to identify the most popular destinations and to compare usage patterns for Internet resources with well-known usage patterns for other types of information.

Online Survey

We were able to do the online survey because the University of Toronto provides Internet access through UTLINK, its online catalog and information system (Data Research Associate's Information Gateway). The university provides the access by connecting library card holders to the Hytelnet software (version 6.3) created by Peter Scott at the University of Saskatchewan. This software automatically makes telnet connections to remote sites and provides the information a user needs to connect to the remote site. There are two menu choices on the UTLINK menu that take users to Hytelnet. One is "Internet Resources," which takes people to sites like Archie servers, e-mail directories and gophers; the

Joy Tillotson is Head, Information Services, Queen Elizabeth II Library, Memorial University of Newfoundland. Her e-mail address is joyt@morgan.uccs.mun.ca. **Joan Cherry** is Associate Professor, Faculty of Information Studies, University of Toronto. Her e-mail address is cherry@fis.utoronto.ca. **Marshall Clinton** is Director, Information Technology Services, University of Toronto Library. His e-mail address is clinton@vax.library.utoronto.ca.

other is "Other Library Catalogues," which takes people to Internet-accessible library catalogs.

We modeled the online survey on a study of remote users of the OPAC at the University of California.⁷ Nineteen multiple-choice and three open questions were presented to one quarter of the users who picked the "Internet Resources" choice on the OPAC menu from December 15, 1993, to January 27, 1994. The questions used in the survey are included in the appendix of this article. The program to run the survey was developed independently of the questions so that it could be used to run other surveys as well. It was written in DIGITAL Command Language, as is the program that allows users to gain access to the OPAC. The survey questions are displayed by a program which reads a text file and waits for a response. Each question is in a separate file which consists of the text of the question and the valid responses. For each response given by a participant, a record is written to a file. The record consists of the response, a question identifier, a participant identifier (generated by the computer) and the date and time. Other programs converted the response file into a format that could be analyzed using SPSS for Windows.

The first eleven questions that were asked, which were presented to users before they began their search, included factual questions about the user's age, sex, and academic status, as well as more subjective questions regarding the user's intentions in using the Internet. Users were then encouraged to go on and use the Internet and, after they had finished, to answer another eleven questions focusing on their satisfaction with the results of their search and with the system.

Using the frequency distributions for the responses to the factual questions, we generated a demographic profile of Internet users and

looked for correlations between demographic variables—for example, between sex and computer ownership or frequency of use of the Internet. The satisfaction ratings provided measures of user reaction to the results of their sessions and to specific aspects of using the Internet via UTLINK. We also examined the data to see if there were significant differences in reaction between subgroups of users based on various demographic variables—for example, age, sex, academic status, or discipline.

Focus Groups

The focus groups were run by two marketing students. There were two groups of undergraduates: one with two female and five male participants and one with two female and eight male participants. Participants were solicited by an open invitation that came up on the screen whenever a user signed on to the OPAC and by the organizers asking for volunteers in selected classes. These methods resulted in groups that were reasonably diverse in academic discipline. The sessions were videotaped. The marketing students who ran the focus groups prepared a report and the first author of this paper reviewed the videotapes. Information was collected from the focus groups about the reasons for using particular Internet sites and about the improvements people would like to see in the system.

Logs of Telnet Connections

The University of Toronto Library keeps logs of outgoing and incoming telnet connections. The logs include the connections made by people choosing "Internet Resources" or "Other Library Catalogues" from the UTLINK menu and a relatively small number of connections made directly from library staff computer accounts. Looking at these logs

made it possible to see general patterns of Internet use as well as to understand some of the reasons for unsuccessful Internet sessions. From the logs, we created a list of the sites to which telnet connections were made from the library computer over a four-month period (September to December 1993) and calculated the number of connections to each site. This enabled us to compare the pattern of Internet use to patterns of use for other information resources, such as those calculated by Richard Trueswell, originator of the 80/20 rule.⁸

One of the survey questions asked people what they found during their Internet session. For those who replied that they found nothing they were looking for, we inspected the log of telnet connections that were made by the user in the time period between answering question number 10 (about their intentions in using the Internet) and question number 12 (whether they found what they were looking for). In this way we were able to identify some causes of unsuccessful searches.

Results

The survey was offered 2,980 times and 505 people responded to at least some of the questions (it was possible to leave the survey after any question) for a "response rate" of 17% (505 divided by 2,980). Of the remaining 2,475 sessions, we received 1,271 replies saying that the user was unwilling to participate and 1,204 saying that the user had already completed the survey. We heard from some people that they had been asked to participate numerous times, so these numbers do not indicate that 1,271 different people were unwilling to participate, since it could have been the same 320 people being asked four times each. We did not have any way of keeping track of users who had already com-

pleted the survey, since no user ID or library card number was recorded in the study. For this reason it is not possible to calculate a response rate that can be compared to manual surveys. In their online survey of remote users of the OPAC at the University of California, Ferl and Millsap had a response rate of 28%, using the same method of calculating "response rate" as we did.⁹ Thus, it seems that in online surveys of this type we are likely to report lower response rates than usually achieved in surveys administered by other means.

Who Is Using the Internet?

Most of the users owned a computer (91%) and a modem (74%) and were using the Internet from home (56%), which was not unexpected. Our results also supported the popularly

held idea¹⁰ that the Internet has more male than female users, since we found that 76% of the users were male. This is from a general university population that is only 48% male.¹¹ In a similar online survey of users of the OPAC in UTLINK conducted in March 1994,¹² the percentage of male users was 51%, so the difference appears to be related to the Internet, not to, say, a greater tendency among men to answer online surveys. The Internet provides access to both academic resources and a wide array of nonacademic sources, whereas OPACs provide access mainly to required academic resources. Therefore, these data might be interpreted as supporting a popular notion of computer usage: that is, that males like to explore the potential of computers, while women prefer to use systems in a task-oriented context.¹³ Figure 1 shows the distri-

bution of Internet users by academic status. Graduate students, faculty, and library staff were present in greater numbers than in the general population at the University of Toronto.

Where Do They Go on the Internet?

Thirteen hundred twenty-five destinations (library catalogs and other resources) were offered to users. Of these, about 1,100 destinations were accessed through the Hynetnet menus by one or more users. From the logs of telnet connections we are able to say which are the most popular destinations for University of Toronto users. Table 1 lists the top ten destinations, which accounted for 33% of the connections. Figure 2 shows the cumulative percentage of use graphed against the cumulative percentage of sites accessed. This graph takes much the same shape as those in Trueswell's article¹⁴ and suggests that Internet use shares the pattern of use he found for journals and books. In this case 80% of the connections were to 13% of the sites. Since the use tends to concentrate on a relatively small number of sites, librarians involved in helping people learn about Internet resources do not need to learn about all of them to be able to help users, but can concentrate on knowing about the popular sites.

The other pattern that was interesting was that use from week to week was consistent among the most popular sites; in other words, sites that were popular in the first week we checked were popular for the entire period, except for slight variations in the order. This pattern was also observed in the popularity rankings kept by various campus information systems (InfoSlug at the University of California Santa Cruz and campus information systems at Texas A&M and Memorial University of Newfoundland). This suggests that if you want to know what

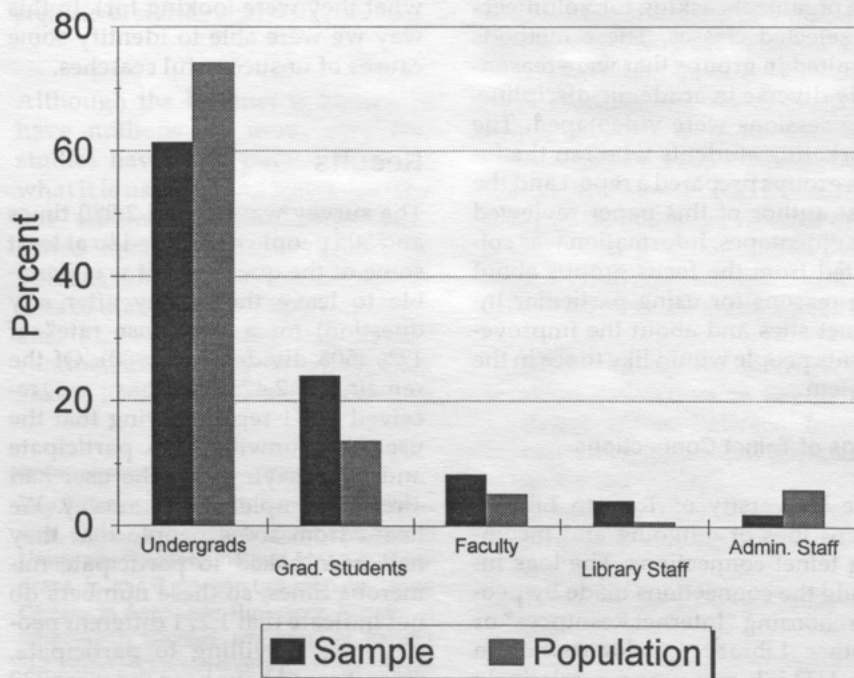


Figure 1
Distribution of users by academic status (N = 497)

Table 1

Ten Internet Sites Accessed Most Frequently (Sept. 1993–Dec. 1993)

1. York University (Toronto, Ont.) library catalog
2. National Capital FreeNet (Ottawa, Ont.)
3. University of Minnesota gopher
4. Cleveland FreeNet
5. University of Iowa Student Computer Association Bulletin Board
6. sun.nsfnet-relay.ac.uk (connection to United Kingdom sites)
7. University of North Carolina Office of Information Technology Bulletin Board
8. Denver FreeNet
9. University of Michigan Library gopher
10. Archie at Advanced Network and Services

the popular destinations are from your campus, only a small sample is necessary.

What we know about why people went to particular sites comes from the focus group interviews and the online survey. One focus group participant said he went to the University of North Carolina site (number seven in the top ten list) because he had found he could get an e-mail address by registering as a user there. He was obviously among those undergraduates at the University of Toronto who were not eligible for free e-mail accounts at the time of this study. It is possible that other users had found that they could get an e-mail address by joining the National Capital FreeNet (number two in the top ten). Other participants spoke of wanting broader access to Internet resources, which could ex-

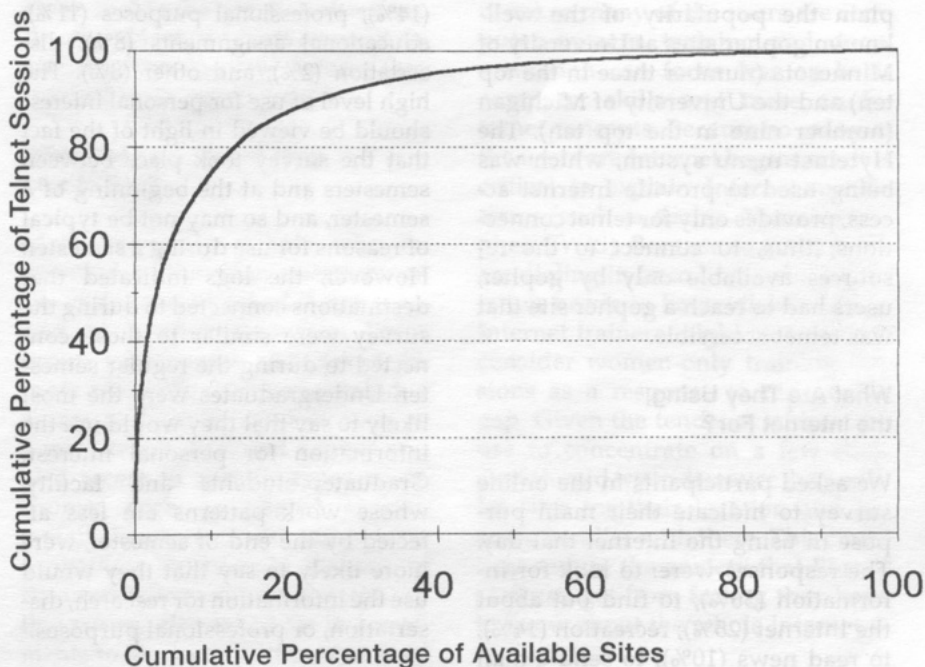


Figure 2

Percentage of use satisfied by percentage of sites accessed (Sept. 1993–Dec. 1993)

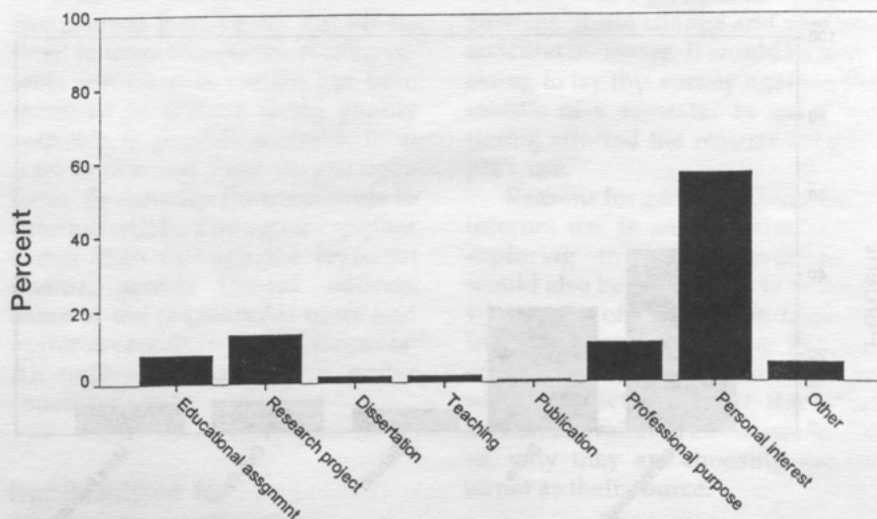


Figure 3

Responses to question "How will you use the information you get from Internet Resources today?" (N = 485)

plain the popularity of the well-known gopher sites at University of Minnesota (number three in the top ten) and the University of Michigan (number nine in the top ten). The Hytelnet menu system, which was being used to provide Internet access, provides only for telnet connections; thus, to connect to the resources available only by gopher, users had to reach a gopher site that was telnet accessible.

What Are They Using the Internet For?

We asked participants in the online survey to indicate their main purpose in using the Internet that day. The responses were: to look for information (36%); to find out about the Internet (25%); recreation (14%); to read news (10%); to send e-mail (8%); and other (7%). Figure 3 shows how users said they would use what they found on the Internet. The majority of users indicated that they would use the information for personal interest (57%). Other uses ranked much lower: research projects

(14%); professional purposes (11%); educational assignments (8%); dissertation (2%); and other (8%). The high level of use for personal interest should be viewed in light of the fact that the survey took place between semesters and at the beginning of a semester, and so may not be typical of reasons for use during a semester. However, the logs indicated that destinations connected to during the survey were similar to those connected to during the regular semester. Undergraduates were the most likely to say that they would use the information for personal interest. Graduate students and faculty, whose work patterns are less affected by the end of semester, were more likely to say that they would use the information for research, dissertation, or professional purposes.

Are They Satisfied with What They Find?

Figure 4 shows what users found in their Internet session compared to what they were looking for. Many people (46%) did not find anything.

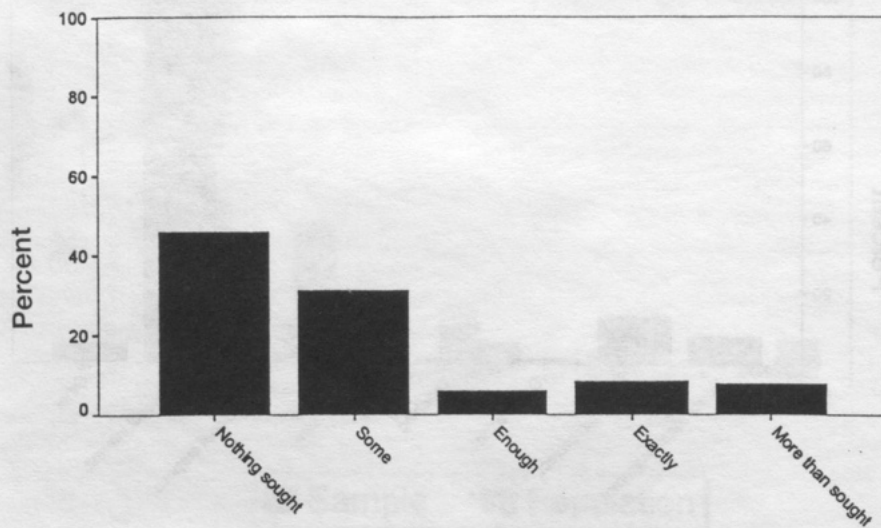


Figure 4 Responses to question "What did you find in your Internet Resources session today?" (N = 328)

The rest found some (31%), enough (6%), exactly (9%), or more (8%) information than they were looking for. We explored the demographic data and telnet logs for things that might explain the large percentage of users who found nothing. We conducted analyses using Chi-square tests to address two questions: (1) Do users who found nothing differ from the rest of the users as a group on any of the demographic variables? (2) Do users who found exactly what they were looking for differ from the rest of the users as a group on any of the demographic variables?

The analyses showed only that first-time or infrequent users were more likely to find nothing than more frequent users (DF [degrees of freedom] = 4, Chi-sq = 13.70324, p [probability] = .00830). The telnet logs showed that many of those who found nothing had failed to make a successful telnet connection. Either they did not choose any of the menu items or they connected to the telnet address several times in quick succession but did not appear to have stayed there long enough to have connected to the database at that address.

Users who said they found exactly what they wanted tended to be daily or weekly users (DF = 4, Chi-sq = 14.53049, p = .00528). Modem owners were more likely to use the Internet daily or weekly rather than less frequently, as were people in computer or library/information science (DF = 44, Chi-sq = 64.65707, p = .02287).

Are They Satisfied with the Way the University of Toronto Library Provides Internet Access?

The scale for satisfaction ratings ranged from 0 (lowest, most negative rating) to 4 (highest, most positive rating). Overall, the rating for general reaction to the UTLINK software which provides access to In-

ternet resources was somewhat positive (2.3). However, F-tests showed that there were significant differences in ratings between groups that found different amounts of information in their Internet sessions ($DF = 4, 298$, $F = 13.4929$, $p = .0000$). Those who found nothing or only some of what they were looking for gave significantly lower ratings than those who found exactly what they were looking for or more than what they were looking for (Duncan's Multiple Range Test; $p = .05$). Similarly, those who found nothing gave lower ratings for ease of learning to use the Internet, while successful users gave higher ratings for ease of learning.

There were also significant differences in ratings for ease of learning between groups based on frequency of use ($DF = 4, 300$, $F = 2.9308$, $p = .0211$) and between groups that were using the Internet for different purposes ($DF = 7, 297$, $F = 2.6510$, $p = .0113$). There were steady increases in ratings for ease of learning as frequency of use increased. Weekly or daily users gave significantly higher ratings (2.2 and 2.4, respectively) than first-time users (1.6) (Duncan's Multiple Range Test, $p = .05$). Those whose purpose was to find out about the Internet gave significantly lower ratings for ease of learning (1.6) than those whose purpose was to read news (2.2), send e-mail (2.3), or look for information (2.3) (Duncan's Multiple Range Test, $p = .05$).

It is not unexpected that greater experience would lead to more success and hence greater satisfaction. It is perhaps more interesting that there were no significant correlations between age, sex, academic discipline, or status and success or ease of learning. We had debated whether to ask people for their age and sex, and compromised by allowing them the option of skipping those two questions. Seven respondents (of 505, or 1.4%) chose to skip the question about age; 22 (or 4.3%)

chose to skip the question about sex. Another 14 respondents stopped answering the survey before they reached these questions.

What Changes Would They Like to See?

Most frequently, people who responded to the open questions on the online survey asked for more access to Internet resources. Focus group participants also asked for more access. First requests included access to Internet services such as e-mail, telnet, FTP, and gopher and also access to specific types of resources such as Usenet groups, e-mail directories, and campus information. Thirty-eight percent of the users who commented in the online survey also asked for improvements to the menu system that gave access to the Internet. Suggested improvements included an index, quicker ways to move through the menus, and subject arrangement or access. Twenty-three percent of those who commented in the online survey asked for more print and online information about the Internet.

Since the survey was conducted, two projects have begun that are related to these comments. A campus-wide information system has been mounted in UTLINK using gopher software to provide access to local information and some remote databases. Eventually, the main route to Internet will be through the gopher rather than through the Hytelnet menus, which should address some of the requests for more and easier access to Internet resources. An online tutorial is also under construction.

Implications for Professional Practice

The results of this study have implications for user training, staff development, and further research. Be-

cause so many of the users are computer owners, training and documentation can focus less on basic computer skills and more on Internet concepts. Because so many of the users are Internet beginners, it is critical to explain Internet-specific terms (such as Archie, Veronica, gopher, CWIS). Because so many users are dialing in from home, documentation needs to be available online. Internet trainers might also want to consider women-only training sessions as a response to the gender gap. Given the tendency of Internet use to concentrate on a few sites, staff should try to discover the popular sites for their location and become familiar with them. This could help reduce the anxiety that library staff can get from feeling they have to know about the whole Internet.

Future Research

In this study, we found a great deal of first-time and exploratory use. However, as the Internet becomes more accessible and more integrated into academic curricula, more research will be required to see if patterns of use change and user satisfaction increases. It would be interesting to try this survey again in the middle of a semester to see if the timing affected the reasons for people's use.

Reasons for gender differences in Internet use is another area worth exploring in future research. It would also be interesting to find out what types of information from the Internet people are using in their educational assignments and research projects; whether the information is available elsewhere and if so, why they are choosing the Internet as their source.

Acknowledgments

We thank staff members in the Infor-

mation Technology Services Office at the University of Toronto who participated in discussions about the wording of the questions used in the survey and helped with the pretest of the survey.

References

1. Hope N. Tillman and Sharyn J. Ladner, "On the Scene: Special Librarians and the INTERNET," *Special Libraries* 83, no. 2 (1992): 127-31.
2. Charles R. McClure, William E. Moen, and Joe Ryan, "Academic Librarians' Perspectives on the Internet/NREN: Key Issues," in *Libraries and the Internet/NREN: Perspectives, Issues and Challenges* (Westport, Conn.: Mecklermedia, 1994), 111-56.
3. Charles R. McClure, William E. Moen, and Joe Ryan, "Public Libraries and the Internet/NREN: New Challenges, New Opportunities" in *Libraries and the Internet/NREN: Perspectives, Issues and Challenges* (Westport, Conn.: Mecklermedia, 1994), 157-98.
4. Michael Eisenberg and Peter Milbury, "The Internet and Library Media Programs: The Technology is Now!," in *Libraries and the Internet/NREN: Perspectives, Issues and Challenges* (Westport, Conn.: Mecklermedia, 1994), 199-220.
5. University of California, Santa Cruz, Communication and Technology Services and the University Library, "InfoSlug Usage Statistics." Computer file (gopher to scilibx.ucsc.edu, path: /About UCSC InfoSlug/InfoSlug Usage Statistics).
6. Texas A&M University, Computing and Information Services, "Gopher Statistics." Computer file (gopher to gopher.tamu.edu, path: /About Gopher (Help & Documentation)/Gopher Statistics).
7. Terry Ellen Ferl and Larry Millsap, "Remote Use of the University of California MELVYL Library System: An Online Survey," *Information Technology and Libraries* 11, no. 3 (1992): 285-303.
8. Richard Trueswell, "Some Behavioral Patterns of Library Users: The 80/20 Rule," *Wilson Library Bulletin* 43, no. 5 (1969): 458-61.
9. Ferl and Millsap.
10. Nancy Kaplan and Eva Farrell, "Weavers of Webs: A Portrait of Young Women on the Net," *Electronic Journal of Virtual Culture* 2, no.3 (1994), computer file (send the message GET KAPLAN V2N3 to list-serv@kentvm.kent.edu).
11. University of Toronto, Office of the Assistant Vice-President (Planning) and University Registrar, *Facts and Figures: University of Toronto: 1992-1993* (Toronto: University of Toronto Press, 1993).
12. Joan Cherry, Joy Tillotson, and Marshall Clinton, "Online Survey of OPAC Users in a Large Academic Library: A Profile of Users and Their Satisfaction," unpublished report, 1994.
13. Barton M. Clark and others, "Gender Gap in the Use of Library Technologies: Evidence, Implications, and Intervention," in *Building on the First Century: Proceedings of the Fifth National Conference of the Association of College and Research Libraries*, ed. J. C. Fennell (Chicago: ALA, 1989), 116-18.
14. Trueswell.

Appendix: Online Survey of Internet Use through the University of Toronto Library (Winter 1994)

Questions A-11 were asked before users started their Internet sessions.

We are conducting a survey of users of Internet Resources to help us improve UTLink. The survey consists of 22 brief questions.

A. Are you willing to participate in the survey?

1. Yes
2. No
3. I have already completed the questionnaire

1. Do you have access to home computer equipment in the house, apartment or residence where you are living?

1. Yes, a computer
2. Yes, a computer and a modem
3. No

2. Where is the terminal you are using today?

1. Your office or work area

2. Home
3. Laboratory
4. Classroom
5. Computer centre
6. College or university library
7. Public library
8. High school library
9. Special library
10. Other

3. Is your terminal:

1. On a University of Toronto campus
2. In Metropolitan Toronto, but not on a University of Toronto campus
3. In Ontario, but outside Metropolitan Toronto
4. In another province of Canada
5. In another country

4. Which of the following describes you best?

1. Undergraduate student (first or second year)
2. Undergraduate student (third or fourth year)

3. Graduate student
 4. Postdoctoral student
 5. Research or teaching faculty
 6. Administrative staff
 7. Research assistant
 8. Library staff
 9. General public
 10. Other
5. Which of these categories describes your main academic field best?
1. Arts and Humanities
 2. Biological Sciences and Agriculture
 3. Social Sciences
 4. Business/Management
 5. Education or Physical Education
 6. Engineering or Applied Sciences
 7. Mathematics or Physical Sciences
 8. Computer Science
 9. Medical/Health Sciences
 10. Law
 11. Library or Information Science
 12. Other
6. Are you a student, faculty, or staff member at or affiliated with the University of Toronto?
1. Yes
 2. No
7. What is your age?
1. 16 or under
 2. 17–24 years
 3. 25–34 years
 4. 35–44 years
 5. 45–54 years
 6. 55–64 years
 7. 65 or over
 8. Skip this question
8. What is your sex?
1. Female
 2. Male
 3. Skip this question
9. How often do you use INTERNET RESOURCES in UTLINK?
1. Daily
 2. Weekly
 3. Monthly
 4. Rarely
 5. Have not used before today

10. What is your main purpose in using INTERNET RESOURCES today?
1. To find out about the Internet
 2. To read news
 3. To read full text of other sources
 4. To send e-mail
 5. For recreation
 6. To look for information
 7. To look for software
 8. Other
11. How will you use the information you get from INTERNET RESOURCES today?
1. For an educational assignment
 2. For a research project
 3. For a dissertation
 4. For teaching
 5. For publication
 6. For professional purposes
 7. For personal interest
 8. Other

Questions 12–22 were asked after users had completed their Internet sessions.

12. What did you find in your INTERNET RESOURCES session today?
1. Nothing you were looking for
 2. Some of what you were looking for
 3. Enough of what you were looking for
 4. Exactly what you were looking for
 5. More than you were looking for
13. General reaction to the UTLINK software which provides access to INTERNET RESOURCES
- | | | | | | |
|----------|-----|-----|-----|-----------|------------|
| (0) | (1) | (2) | (3) | (4) | (x) |
| terrible | | | | wonderful | no opinion |
14. Learning to use the Internet
- | | | | | | |
|-----------|-----|-----|-----|------|------------|
| (0) | (1) | (2) | (3) | (4) | (x) |
| difficult | | | | easy | no opinion |
15. Destinations or sources can be accessed in a straightforward manner
- | | | | | | |
|-------|-----|-----|-----|--------|------------|
| (0) | (1) | (2) | (3) | (4) | (x) |
| never | | | | always | no opinion |

16. Choices on UTLink menus for INTERNET RESOURCES

(0)	(1)	(2)	(3)	(4)	(x)
confusing				clear	no opinion

17. Number of terminals provided by the library

(0)	(1)	(2)	(3)	(4)	(x)
not enough				enough	no opinion

18. Location of terminals provided by the library

(0)	(1)	(2)	(3)	(4)	(x)
inconvenient				convenient	no opinion

19. Dial-in access at times convenient for me

(0)	(1)	(2)	(3)	(4)	(x)
usually unavailable				usually available	no opinion

20. What would you like to see added to INTERNET RESOURCES in UTLink?

21. Do you have any suggestions for reorganizing the UTLink menus for INTERNET RESOURCES?

22. Do you have any more comments about INTERNET RESOURCES in UTLink?

Index to Advertisers

ALA Editions	200
Ameritech Library	cover 2
Axxess Information	133
Blackwell North America	130
Gaylord	129
LITA	172
Library Technologies	160
Minolta	cover 3
Online, Inc.	134
Todd Enterprises	cover 4
WLN	154

Letters and Announcement

Dear Editor:

Thank you for publishing the fascinating article, "Digitization and the Creation of Virtual Libraries, The Princeton University Image Card Catalog—Reaping the Benefits of Imaging," by Eileen Henthorne (*Information Technology and Libraries*, March 1995).

A tiny quibble. The author refers to 3-by-5-inch catalog cards. Melvil Dewey, in addition to establishing the "Library Bureau" to sell library equipment and supplies, established the less successful "Metric Bureau."

Because of Dewey's interest in the metric system our catalog cards are "7.5cm-by-12.5cm."

During the depression years, to save money, many 3-by-5-inch cards were used and were interfiled in the public card catalog of the Orange (N.J.) Public Library. As 3-by-5-inch cards are slightly larger, one could easily see the frayed dirty top edges of these cards.—Marvin H. Scilken, Editor, *The Unabashed Librarian*

Dear Editor:

The March 1995 article comparing access to dissertations via WorldCat (the OCLC Online Union Catalog) on FirstSearch and *Dissertation Abstracts* on CitaDel (Stephen Perry and Lutishoor Gupta, "Access to Information in Both CitaDel and

FirstSearch: A Comparative Study of Dissertation Coverage," *Information Technology and Libraries* 14, no. 1 [March 1995]: 17–29) pointed out important differences in coverage of academic literature in these databases. In January 1995, OCLC was delighted to announce *Dissertation Abstracts'* availability on FirstSearch, via subscription and per search. In addition, FirstSearch includes more than fifty other databases, and several of these (e.g., *IN-SPEC*, *MLA Bibliography*) include coverage of relevant dissertations.—Tam Dalrymple, Manager, Database Acquisitions & Product Management Section, Reference Services Division, OCLC

Crawford Wins LITA/Library Hi Tech Award

Walt Crawford, senior analyst for the Research Libraries Group, is the 1995 recipient of the LITA/Library Hi Tech Award. The award, \$1,000 and a plaque donated by Pierian Press, is given to an individual or institution for work that shows outstanding communication for continuing education in library and information technology. The award was presented at the LITA/ACRL Joint President's Program during the ALA Annual Conference in Chicago.

"The award committee was unanimous in the selection of Walt Crawford to be the recipient for this award," said Michele Dalehite, chair of the award committee. "Through

his twelve books and sixty-four articles, Walt has contributed significantly to the education, awareness, and edification of his colleagues in a profession that he has chosen to adopt as his own."

The committee specifically recognized Crawford for several publications: *MARC for Library Use: Understanding Integrated MARC* (1989), *Technical Standards: An Introduction for Librarians* (1991), *The Online Catalog Book: Essays and Examples* (1992), and *Future Libraries: Dreams, Madness and Reality*, with Michael Gorman (1995).

Crawford has held several positions at the Research Libraries Group, including programmer/ana-

lyst and manager of the production batch group. He was also programmer/analyst at the University of California, Berkeley.

A member of the *ITAL* editorial board, Crawford was editor of the *LITA Newsletter* from 1985 to 1994. He is past president (1992–93) of LITA and served on the LITA board (1988–91). Crawford was a member of the LITA Technical Standards for Library Automation (TESLA) Committee (1978–82) and LITA Machine Readable Bibliographic Instruction (MARBI) Committee (1985–87). He also served as a member of the Association for Library Collections and Technical Services (ALCTS) Free MARC Committee (1990–91).



Your Index to the World of Research ~ Now available in both print and CD-ROM!

Guide to Reference Books, 11th Edition

Robert Balay, editor

ALA Editions proudly announces the Eleventh Edition of one of the most powerful and authoritative reference tools ever developed for librarians and researchers. Familiar to most as the master source from which reference tools are identified, *Guide to Reference Books, 11th Edition*, maintains a standard of quality against which other guides are measured.

With nearly 16,000 entries, this landmark edition organizes and describes the most important reference works for today's research and scholarship, including carefully selected items from every part of the world.

GTRB is everything the Internet is not: Organized, Indexed, Advisory & Authoritative.

Features of the 11th Edition include:

- Abundance of electronic resources—CD-ROMs, online databases, and more
- Annotations of approximately 16,000 reference/research sources
- Entries describe special features and compare similar or competing titles
- Scholarly research materials as well as general reference sources needed in research libraries
- More extensive guidance on use of the reference works, with expert opinions on their merits and shortcomings

The CD-ROM version saves you valuable time. No more manual searching or photocopying! Quickly compile special-purpose bibliographies! Contains all of the authoritative information in the print version but adds benefits you can only get through today's hottest format, such as:

- Instant hypertext jumps between cross-references
- Fill-in-the-blank query templates for fast, easy searching
- Search terms across every field, including annotations, to locate all sources related to a key term in just one try
- Save searches into custom-made, topical bibliographies that can be printed on demand
- Create your own pop-up notes for each title.

FOLIO. System Requirements: IBM-compatible 386 or higher (recommended: 486SX33 or higher); at least 4 MB of available RAM (8 MB recommended); Windows 3.1 or later; 6 MB of free space on your hard disk.

The print version is ideal for multiple users and browsing, providing access through its familiar topical arrangement and its comprehensive, 400-plus page, single-alphabet index. Sewn-case binding offers years of authoritative service.

More than 50 distinguished reference librarians practicing in research libraries from Harvard to the University of California prepared this wealth of essential reference information. Editor Robert Balay is Reference Editor, *Choice*, and former Head of Reference, Yale University.

\$275.00 hard-cover ■ 2,000p. ■ 1996 ■ ALA Order Code 0669-9-2028

\$475.00 CD-ROM ■ 1996 ■ ALA Order Code 0675-3-2028

SPECIAL SAVINGS—

Get both the CD-ROM and print version together for only \$650.00
ALA Order Code 0676-1-2028



TO ORDER, CALL 800-545-2433 AND PRESS 7



American Library Association
Book Order Fulfillment
155 N. Wacker Drive • Chicago IL 60606

Building Information Literacy Using High Technology

A Guide for Schools and Libraries

By Roxanne Mendrinos. Englewood, Colo.: Libraries Unlimited. 1994. 190p. \$25 (ISBN 1-56308-032-X).

The results of two research studies and the author's experience as a library media specialist serve as the basis for this valuable reference guide for those in a school library setting who are facing the challenge of the various new technologies such as CD-ROM, online telecommunications, and the Internet. Mendrinos focuses on "the process of information literacy in a resource-based learning environment" as the most appropriate way to prepare students for their futures. The new technologies enhance this process, particularly when school library media specialists and teachers collaborate on instructional projects.

Half of the book describes ways to collaborate using high technology. Four case studies are provided from Archbishop Carroll High School, Radnor, Pa.; Nokomis Regional High School, Newport, Maine; Queen Anne's County School, Centerville, Md.; and Thurston Middle School, Westwood, Mass. One chapter is devoted to developing information literacy in the classroom; another includes detailed lesson plans for five units. This part of the book will be most helpful to those striving to incorporate high technology into the curriculum. The first half of the book provides technical information about online databases, modems, software, vendors, networks, telecommunications, the Internet, and CD-ROM as these relate to the educational setting. One chapter is devoted entirely to CD-ROM and the

networking environment, including ten illustrations of the various aspects of networks. Appendixes include an epilogue: "Developing a Critical Bibliography," a glossary, references, and an index.

Those readers new to technology will find that this book satisfactorily answers many technological questions, but its greater significance will be its modeling of technology use within the curriculum through library media specialist and teacher collaboration.—*Janet L. Barr, Eastern Suffolk BOCES School Library System*

Hands-On Mosaic

A Tutorial for Windows Users

By David Sachs and Henry Stair. Englewood Cliffs, N.J.: Prentice-Hall, 1995. 366p. Paper, \$29.95 (includes diskette) (ISBN 0-13-17231-9).

This is a disappointing book, not because it does not tutor but because it deals with details that will frustrate many Internet beginners. The surge of interest in the World Wide Web has spurred the development of software packages that make it easy for anyone to connect to the Internet without having to contend with IP (Internet Protocol) addresses and SLIP (Serial Line Internet Protocol) connections. For example, I installed NetCruiser (from NetCom) and was browsing Web pages from Disney in just under ten minutes, and that included registering myself as an account holder.

In contrast, the authors walk me through many steps in the three parts of the book, including the following: Part I: (1) Install the Chameleon sampler disk (included with the book, courtesy of NetManage, Inc.); (2) initiate a SLIP connection with an Internet service provider; (3) telnet to the InterNIC gopher site. Part II: (4) FTP to one site to obtain PKWare

software; (5) FTP to the National Center for Supercomputing Applications (NCSA) to obtain a copy of Mosaic ver. 1 (the sixteen-bit version); (6) unzip and install Mosaic. Part III: (7) Use Mosaic to browse resources at various sites.

By now, much time has elapsed and I am on page 149. I have learned how to install and use some software, but I am about to discover that I probably need more software—the thirty-two-bit version of Mosaic as well as helper applications such as LView for viewing graphics. All this has certainly been hands-on and not dry, theoretical instruction. However, if all I wanted was to look at a popular Web site, it's been too much work. And if I wanted some instruction on using Mosaic, other books (see below) provide much more depth than this volume.

To the authors' credit, the steps are clearly illustrated with screen shots throughout the book, and tips are provided to alert the reader/user to special software features.

Computer, networking, and Internet terminology is crisply explained in the text where appropriate. The authors also provide, in the appendix, three lists of Internet service providers (InterNIC, NIXPub, and PDIAL lists).

This book is designated by the publisher as being appropriate for beginner and intermediate users, but I would advise beginners to consider subscribing to an Internet service provider such as Pipeline or Netcom instead. Both of these companies will provide graphical Web browser software that does not require a SLIP or PPP connection. It won't be Mosaic, but it will work in minutes instead of hours. And for a better book on the subject, consider Paul Gilster's *Mosaic Navigator* or the *Open Computing Guide to Mosaic* by Levi Reiss and Joseph Radin.—*Ka-Neng Au, Rutgers University, Newark, New Jersey*

HTML for Fun and Profit

By Mary E. S. Morris. Mountain View, Calif.: Sun Soft Press (Prentice), 1995. 264p. \$35.95 (includes CD-ROM) (ISBN 0-13-359290-1).

Mary E. S. Morris has written an introductory guide to both creating documents in HTML (hypertext markup language) and setting up a server for such documents. The accompanying CD-ROM, which contains programs, tools, and sample HTML and Perl scripts, can be installed on UNIX (Solaris 1.x/Sun OS 4.x; Solaris 2.x/Sun OS 5.x), Microsoft Windows NT (not Windows 3.1), and Macintosh systems. Most valuable for HTML authors are the sections on forms, clickable images, and tables. A brief chapter on filters to convert preexisting files and authoring tools directs users to the Internet for more information, though the CD-ROM includes a Macintosh filer (RTF or Rich Text Format) and authoring tools for Mac and Windows (HTML Assistant and HoTMetaL) systems. The publication may expedite the process of setting up a server by including in one place files that have been downloaded from the Internet together with directions for installing them. However, supplementary texts on servers for a particular platform and Perl programming are recommended, as well as updating by FTP some CD-ROM files from mid-1994.

Full installation of the files on the CD-ROM begins by creating an HTTP server on one of the named platforms. Material on the implications of HTML for the configuration and functioning of the server makes up a significant part of the book. An alternative used by the reviewer was to copy and install on an established HTTP server those CD-ROM files necessary to facilitate interactive exercises (original chapter assignments and executable files). Most of the digital material is available in the public domain, though Morris cau-

tions readers to note information with each piece of software.

HTML authors can use the printed book without the digital files, since Morris provides her original illustrated HTML files. Even though these files are well described in the text, the effects of actions upon the files lose their immediacy and sense of fun usually possible only through actions on executable files accessed via a browser through the Common Gateway Interface (CGI).

Readers may object to the organization of the book as a fixed tutorial with step-by-step directions and with interactivity limited to display on the browser. Brief concepts are joined in a "cut-and-paste" manner, including irritating repetition and perorations on such subjects as emotags. There are better books on the market for a concise introduction to the easy HTML language, but none that I know with guides to forms and tables and the dual goal of educating both the HTTP server administrator and HTML author.

The design of the pages looks like HTML text with many illustrations: no centering or running heads, notes inserted between rules in the text. The CD-ROM merely stores the collection of ASCII files in directories and subdirectories. The complex text has few errors, and there is an extensive glossary. An appendix extends the usual bibliography of books to include FTP sites, news groups, and Web addresses. The publication is valuable for this brief window of time.—*Mary Kay Duggan, University of California, Berkeley*

The HTML Sourcebook

By Ian S. Graham. New York: Wiley, 1995. 416p. \$29.95 (ISBN 0-471-11849-4).

Who among us remains unaware of the exponential growth of the World Wide Web? In April 1993, usage of the Internet for WWW traffic con-

sumed a mere .1 percent of the bandwidth (as measured in bytes), and in that year Hyper Text Transfer Protocol (HTTP) was the twenty-eighth most used protocol. In April 1994, usage of the Internet for WWW traffic had risen to consume 4.7 percent of the bandwidth, and HTTP was the sixth most used protocol. In April 1995, usage for WWW traffic climbed to consume more than 26 percent of the bandwidth, and HTTP became the single most used protocol on the Internet, edging out file transfer protocol (FTP). Accordingly, two years ago we had very few hypertext markup language (HTML) editors from which to choose, but now we have many. We also had little need to become fluent in HTML, but now most of us feel a great need to become so. This book should provide the reasonably experienced user of the WWW a good measure of that needed fluency.

Ian Graham clearly knows his subject and employs a sprightly style that probably renders this sort of subject matter as approachable as can be done, which is only to say that this is a practical book filled with much detailed information on the topic. Chapter 1 is a general introduction to HTML and to design issues that surround its use. Chapters 2, 3, and 4 make up the technical core of the book. Here Graham covers the details of HTML commands, syntax, and the rules for the construction of URLs (Uniform Resource Locators). He also covers the interaction between WWW clients and HTTP servers, and the use of the Common Gateway Interface (CGI) specification. In chapters 5 and 6 Graham provides additional tools related to converting documents to HTML and includes directions for obtaining these tools on the Internet. Of particular interest are the specialized tools to convert files written in Rich Text Format (RTF), LaTeX, and others. In chapter 7 Graham reviews Web browsing software available at

the time the book was written. Unfortunately, the number of browsers is growing so fast that no book can cover them adequately. In chapters 8 and 9 Graham covers issues involved in setting up an HTTP server and reviews some common server packages. Finally, Graham provides some examples of interesting WWW sites and comments from their creators. In all of this Graham pays very even-handed attention to all of the environments in which designers can implement browsers and servers: MS-DOS, Windows, OS/2, the Macintosh, UNIX and VAX/VMS, NeXt, and the Amiga.

The book also includes several appendixes: appendix A covers the use of the ISO Latin-1 character set; B, the use of Multipurpose Internet Mail Extensions (MIME); and C, the use of Archie clients to search for appropriate software. The use of character sets on the Internet is a particularly timely topic, and one that Graham might have covered more deeply. MIME too is a hot topic, but with respect to HTML and the WWW is probably adequately covered here. A short glossary and an index complete the book.

The organization of material serves the subject well. Particularly in the core chapters that cover commands, the explanations, descriptions, and examples fit well into the scheme. My one criticism is that the layout is confusing. The use of bold-face type for some headings, different font sizes for others, and horizontal rules of different weights to set off others does not help the reader focus on the material but is slightly distracting and confusing. In summary, this is a very practical and useful book about one of the fastest growing and changing fields in librarianship. As such it is necessarily a little dated, particularly when discussing different commercial browsers. However, the book is well worth acquiring and reading.—*Fritz Schwartz, The Faxon Company*

Implementing an Automated Circulation System

By Kathleen G. Fouty. *A How-To-Do-It Manual*. New York: Neal-Schuman, 1994. 220p. Paper, \$39.95 (ISBN 1-55570-175-2).

This title is a fine addition to a practical and insightful series. Any library implementing a new library system is faced with a very big challenge. For a circulation department this challenge is even greater. Constantly on the front lines, the department's computer system can determine a patron's perception of the rest of the library. Introducing a new system to this area must be done very carefully.

Fouty covers the topic of implementing a new circulation system very well. The chapters follow a logical progression from naming a project manager to assessing the installed system. In between, the chapters cover installing software and hardware, converting data, establishing procedures and policies, developing manuals and a staff training program, facilitating staff adjustment to the new system, working up the system. The author states that the book's concepts are "fundamental to most automated circulation implementation projects, and can be adapted readily for application in academic, public, school or special libraries" (p. ix). The reader will find the author true to her word. Specific circulation systems can be designed for a particular type of library, but the administrative concerns can be the same for libraries of any size and type. This book provides an excellent review of things administrators and project coordinators need to keep in mind during the implementation process. In addition, many of the topics covered could easily be applied to other modules of an integrated system; the title for this book

could easily have been *Implementing an Integrated Library System*.

The book is full of checklists and guidelines and includes wide margins that would allow it to be used as a workbook. The headings are well laid out, and all topics are thoroughly covered. The author's recommendations are sound and obviously based on experience. Chapters 10 and 11 are particularly strong in their coverage of staff adjustment and training, with a sample curriculum and outlines. The book's only weakness is the timeliness of the references: most of them date back to the early mid-1980s. However, since specific circulation products are not discussed and the author concentrates on administrative issues that are still relevant in the 1990s, this is not a serious flaw. At \$39.95 this title is a good value and a worthwhile purchase for any library considering implementing a new integrated or circulation system.—*Susan E. Cleyle, Memorial University of Newfoundland*

Introduction to Reference Sources in the Health Sciences

By Fred W. Roper and Jo Anne Boorkman. 3d edition. Metuchen, N.J.: Medical Library Association and Scarecrow, 1994. 301p. \$35 (ISBN 0-8108-2889-8).

Reference and Information Services in Health Sciences Libraries

Ed. by M. Sandra Wood. Vol. 1 of *Current Practice in Health Sciences Librarianship*. Metuchen, N.J.: Medical Library Association and Scarecrow, 1994. 371p. \$39.50 (ISBN 0-8108-2765-4).

"The purpose of reference services," states the introduction to *Reference and Information Services in Health Sciences Libraries*, "is to assist users in securing information and in using

library resources" (p. xvii). These two works are essential guides toward fulfilling this mission, one focusing on the resources needed to perform this task, the other on the skills and practices.

It has been ten years since the second edition of *Introduction to Reference Sources in the Health Sciences* was published, and this third edition marks a welcome revision. Like the earlier editions, this edition emphasizes the basic reference tools in health sciences and is directed toward meeting the needs of all health sciences libraries, from the small hospital to the large academic medical center.

Changes since the previous edition reflect the changes that have occurred in health sciences librarianship in the interim. The chapter titled "Online Bibliographic Databases" has been renamed "Electronic Bibliographic Databases" to reflect the new direction in reference sources, and the chapter on audiovisual reference sources has been expanded to include microcomputer and multimedia reference sources.

The book is divided into three sections. It begins with a thorough discussion on the selection, organization, and maintenance of the reference collection, including electronic reference sources. The second section focuses on bibliographic resources. It documents sources of information for monographs, periodicals, government reports, technical documents, conferences, and translations, as well as those sources that provide access to the literature: the indexing and abstracting services and electronic bibliographic databases. The final section of the book reviews information sources. It describes dictionaries, handbooks, directories, and biographical sources, as well as the more specialized information resources on drugs, health statistics, and grants.

One of the strengths of this work is its historical dimension. The chap-

ter on indexing and abstracting services, for example, presents an engrossing discussion of the development of the National Library of Medicine's *Index Medicus*, with a detailed explanation of the MeSH (Medical Subject Headings) structure that should be required reading for any health sciences librarian. Moreover, there has been a sincere effort to integrate electronic reference sources into the work, and in most chapters the discussion moves easily between electronic and print sources of information. The one exception to this is the discussion of sources that provide access to the literature, where one chapter reviews print sources and a separate chapter examines electronic databases. To this reader, the separation is awkward, and it is hoped that future editions of the work will be able to integrate these reference media more fully.

Reference and Information Services in Health Sciences Libraries likewise represents a revision of a standard reference work. The work constitutes the first volume in the Current Practice in Health Sciences Librarianship series, a revision of the 1982 *Handbook of Medical Library Practice*. If this work is any indication, reference services have come a long way during this time. Whereas the previous edition offered only two chapters on reference in the volume on public services, this edition devotes an entire volume to the topic.

The work is divided into six chapters, each covering specific aspects of reference services in health sciences libraries. It opens with a description of traditional reference services, with emphasis on the reference interview. The second chapter, on databases and database searching, is the most lengthy and in many ways forms the core of this work. This chapter provides a comprehensive review of online services in health sciences libraries. Encompassing both mediated and end-user

searching, the chapter outlines the historical development of online services and examines the organization and management of search services. Information provided is of immediate practical application, with tables of bibliographic databases, an inventory of file management programs, lists of vendors, criteria for evaluating search services, and a review of search techniques and procedures.

Subsequent chapters investigate microcomputers in reference service; the management of reference services; and specialized reference services, such as clinical librarianship, consumer health services, and fee-based services. The final chapter summarizes current technologies and projects future directions. Internet services, document distribution, decision-support systems, and electronic publishing are presented as examples of new developments that will influence how reference services are provided in health sciences libraries in the next decades.

Throughout the work, the emphasis is on emerging trends in health sciences reference services and the changing role of the reference librarian. The keystone is technology. The new reference librarian must be technologically literate; she or he must have basic knowledge about computers, databases, and telecommunications systems, as well as print resources.

Reviewing these two works at the same time has made for an interesting assignment. At first glance, *Reference and Information Services in Health Sciences Libraries* seems the more exciting, forward-looking work, whereas *Introduction to Reference Sources in the Health Sciences* appears old-fashioned. This initial impression, however, is erroneous. The two works are complementary, each with its own function, and each an equally important tool. Both works belong in any library school class on health sciences librarianship and

should be on the shelf of practicing librarians in medical libraries of all types.—*Susan Jacobson, Columbia University*

Privatizing Government Information

The Effects of Policy on Access to Landsat Satellite Data

By Kathleen Eisenbeis. Metuchen, N.J.: Scarecrow, 1995. 327p. \$42.50 (ISBN 0-8108-2934-7).

Land remote sensing from satellites began in the late 1960s with development of the Earth Resources Technology Satellite (ERTS), first launched in 1972. This satellite, renamed Landsat 1, and later versions Landsats 2 through 5, have since provided over twenty years of earth observation data to a broad user community including the agricultural community, global change researchers, state and local governments, commercial users, and the military. Privatization of this data-gathering program was discussed first in the Carter administration, when the system was transferred from the Department of Defense to the National Oceanic and Atmospheric Administration, and became a reality in the Reagan Administration with passage of the Land Remote-Sensing Commercialization Act of 1984. From the early misgivings voiced at the program's inception, through the politics and policies leading up to and culminating in the passage of the 1984 law, the Landsat privatization process and its subsequent effects on Landsat data users are examined in this new work by Kathleen Eisenbeis, a frequent writer and speaker on information access and governmental policies.

Placing the Landsat commercialization in the contexts of both public policy and information pol-

icy, Eisenbeis asks and answers the basic question, "Does privatization affect access to government information?" Through numerous research methods, including content analysis of Landsat research in remote sensing journals and a survey of data users, she is able to measure the negative effects of this particular policy on the ability of academic geographers to conduct remote sensing research and to instruct students in the use of Landsat data.

Eisenbeis also traces Landsat's recent path back to the federal government, a process hastened by such diverse influences as the growth of GIS technology, the Gulf War, and the desire of Congress to "avoid paying high prices in the future" for what is essentially the federal government's own data. The 1984 law was repealed in 1992, and since this book went to press, the Clinton administration has outlined a new policy (Presidential Decision Directive/NSTC-3, 5 May 1994) that "provides for the continuance of the Landsat . . . program, assures continuity of Landsat-type and quality of data, and reduces the risk of a data gap."

Eisenbeis' book will be of interest to a broad spectrum of readers, including those who are generally unfamiliar with Landsat and/or remote sensing. The detailed history of the privatization process is a fascinating look into political decision-making processes. With the inclusion of lengthy chapters on government information policy concepts and a follow-up literature review, a legislative history of the 1984 act, and the painfully honest questionnaire responses on the effects of Landsat's commercialization on the academic community, the book will also appeal to public policy researchers, librarians, and others in the field of information policy. In this time of increased Congressional pressure for privatization of government programs, we would do well to increase our knowledge of past ex-

periences with this particular method of government downsizing.—*Mary McInroy, University of Iowa*

Reference Services in the Humanities

Ed. Judy Reynolds. Binghamton, N.Y.: Haworth 1994. 214p. \$39.95 (ISBN 1-56024-692.8). Also published as *The Reference Librarian*, no. 47 (1994).

This is a selective compilation of fourteen articles covering various aspects of supplying reference services in the humanities. It is organized into four groups: research questions and challenges in selected disciplines, descriptions from the field, political issues in the humanities, and theories and ideas for the future. Subject areas covered include crafts, motion pictures, music, modern languages, history, American Indians, censorship, special collections, and theater.

The book covers a wide range of reference services and programs developed to respond to the needs of patrons engaged in humanities research. In her introduction, the editor points out that humanities researchers delight in harmonies of meaning and use metaphoric language. The precision of science research and the semiprecision of social science research are lacking. Reference librarians must be aware of the metaphoric and imprecise nature of many patron queries and of the limitations of language reflected in both subject headings and titles. The increasing dependence on online catalogs, with their ability to manipulate language, offers advantages over traditional card catalogs and printed indexes that mature humanities researchers often fail to exploit. This situation is compounded by the tendency of such researchers to do their research in isolation and not seek the advice of reference librarians.

Newer humanities researchers are more apt to use new technologies for research but still often avoid seeking reference assistance.

The fourteen different contributions adopt a conceptual approach to various facets of offering reference service in the humanities to a wide variety of patrons. They make interesting reading and are thankfully free of the avalanche of statistics on patterns of patron behavior that abound in library literature. My choice for the best of the lot is Scott Stebelman's "Vocabulary Control and the Humanities: A Case Study of the MLA International Bibliography," which he describes as a prototype for successful contextual searching. Readers might favor other contributions based on personal interest. A sobering note is struck by Louis Charles Willard in his concluding piece, "The Library Yet to Come," in which he foresees the possibility that newly trained scholars will become more sophisticated acquirers of the trappings of scholarship but less thoughtful, less reflective practitioners. He exhorts librarians to shed their introspective, passive character and to exploit better the resources of the electronic world. This collection is a solid if not profound contribution to the literature on reference service in the humanities.—*Thomas A. Bourke, The New York Public Library*

TCP/IP for the Internet The Complete Buyer's Guide to Micro-Based TCP/IP Software

By Marshall Breeding. Westport, Conn.: Mecklermedia, 1995. 305p. Paper, \$24.95 (ISBN 0-88736-980-4)

How does a library selector handle a book with a very short shelf life, one covering a hot technology topic and oriented to a select group of people experienced enough to understand

the subject? Breeding's *TCP/IP for the Internet* fits solidly into this category. Breeding acknowledges the short shelf life of this book when he says, "the demand for TCP/IP stacks bundled with network applications should last for as much as another two years" (p. 20).

Breeding does an admirable job in doing what he sets out to do, which is to explain and review the primary competing TCP/IP products for connecting to the Internet. The first chapter offers basic information and concepts regarding TCP/IP protocols and applications. These explanations are necessary to understand the product reviews and make informed purchasing decisions. The book is filled with illustrations and charts to help represent his points.

Chapter 2 describes issues and concerns for implementing TCP/IP in a Windows environment. Breeding bases his reviews on how the vendors provide network interface drivers that access the network hardware, implementation of the TCP/IP protocols, and Internet applications such as FTP (file transfer protocol) and telnet. This chapter also gets quite technical as he describes the differences between ODI (Open Data Interface), NDIS (Network Device Interface Specification) and packet drivers, and also TSR (Terminate and Stay Resident), VxD (Virtual Device Drive) protocol and .DDL (Dynamic Data Link) TCP/IP kernels. It also illustrates that this topic is not for the novice computer user.

Chapters 3 through 18 review fifteen vendors' products. In these reviews, Breeding provides postal and e-mail addresses, corporate history, general product descriptions, product summary checklists, and pricing information. On a more detailed level he describes how each product handles the TCP/IP stack, installation, memory management, network performance, application utilities, user

interface, and documentation. Each review ends with a Buyer's Guide Report Card that gives an excellent overview of the product based on price, value, design, performance, and applications. These report cards may be the primary benefit most readers will derive from this book.

Chapter 19 compares the various products, primarily through the use of charts. Breeding makes no recommendations, but lets the reviews stand on their own.

Chapter 20 provides the same information found in chapter 2, but oriented to the Macintosh environment. Chapters 21–24 review Macintosh products along the same lines used for reviewing Windows products. The book ends with an excellent glossary and a good index.

Each library will have to decide how to handle specialized, short term, technology-related books. If your policy is to buy these books and make them available to the public, then this is a fine example of a highly specialized title. *TCP/IP for the Internet* may also be of interest to larger libraries who have network administrators.—*Valerie J. Horton, New Mexico State University Library*

Other Recent Receipts

Access, Ownership, and Resource Sharing. Ed. Sul H. Lee. New York: Haworth, 1995. 125p. \$24.95 (ISBN 1-56024-727-4).

Advances in Collection Development and Resource Management. Vol. 1, 1995. Ed. Thomas W. Leonhardt. Greenwich, Conn.: JAI, 1995. 186p. \$73.25 (ISBN 1-55938-213-9).

ASIS Thesaurus of Information Science and Librarianship. Ed. Jessica L. Milstead. Medford, N.J.: Learned Information, 1994. 139p. Paper, \$39.95 (ISBN 0-938734-80-6).

Automation for School Libraries: How To Do It From Those Who Have Done It. Ed. Teresa Thurman Day, Bruce Flanders, and Gregory Zuck. Chicago, Ill.: American Library Assn.,

1994. 138p. Paper, \$18 (ISBN 0-8389-0637-0).
- Barclay, John. *The Seventy Year Ebb and Flow of Chinese Library and Information Services, May 4, 1919, to the Late 1980s*. Metuchen, N.J.: Scarecrow, 1995. 351p. \$42.50 (ISBN 0-8108-271-1).
- Building Partnerships: Computing and Library Professionals*. Proceedings of Library Solutions Institute Number 3, Chicago, Illinois, May 12-14, 1994. Berkeley, Calif.: Library Solutions Press, 1995. 102p. Paper, \$29 (ISBN 1-882208-18-8).
- Business Resources on the Internet Plus: A Hands-on Workshop*. By Gary R. Peete. Includes Windows and Macintosh diskettes of presentation slides. Berkeley, Calif.: Library Solutions Press, 1995. 252p. Paper, \$60 (ISBN 1-882208-110).
- Cataloging Service Bulletin Index: An Index to the Cataloging Service Bulletin of the Library of Congress*. Nos. 1-66, Summer 1978-Fall 1994. Comp. Nancy B. Olson. New York: Soldier Creek, 1995. 89p. Paper, \$35 (ISBN 0-936996-66-8).
- CD-ROMs in Print: An International Guide to: CD-ROM, CD-1, 3DO, MMCD, CD32, Multimedia & Electronic Book Products* (includes CD-ROM). Ed. Regina Rega. Westport, Conn.: Mecklermedia, 1995. 1,117p. Paper, \$125 (ISBN 0-88736-983-9).
- Comer, Douglas E. *The Internet Book: Everything You Need To Know About Computer Networking and How the Internet Works*. Englewood, N.J.: Simon and Schuster, 1995. 310p. Paper, \$24.95 (ISBN 0-13-151565-9).
- Comer, Douglas E., and David L. Stevens. *Internetworking with TCP/IP, Volume II: Design, Implementation, and Internals*. Englewood Cliffs, N.J.: Prentice-Hall, 1994. 597p. \$60 (ISBN 0-13-125527-4).
- Copyright Theft*. By John Gurnsey. Brookfield, Vt.: Ashgate, 1995. 196p. \$48.95 (ISBN 0-566-07631-4).
- Current Awareness, Current Techniques*. By Feona Hamilton. Brookfield, Vt.: Ashgate, 1995. 175p. \$49.95 (ISBN 0-566-07626-8).
- Directory of European Research and Development*. 2 vols. New Providence, N.J.: Bowker-Saur, 1995. 1085p. Paper, \$315. Offered in combination with *Who's Who in European Research and Development 1995* at \$400/set. (ISBN 1-85739-092-X).
- Disaster Planning for Library and Information Services*. Ed. John Ashman. London: Aslib, 1995. 51p. Paper, \$25, \$20 members (no ISBN).
- Disaster Prevention and Response for Special Libraries: An Information Kit*. By Miriam Kahn. Washington, D.C.: Special Libraries Assn., 1995. Paper, \$20, \$15 members (ISBN 0-87111-436-4).
- Distance Education: An Annotated Bibliography*. By Terry Ann Mood. Englewood, Colo.: Libraries Unlimited, 1995. 200p. Paper, \$27 (ISBN 1-56308-160-1).
- Doyle, Christina S. *Information Literacy in an Information Society: A Concept for the Information Age, June 1994*. New York: ERIC Clearinghouse on Information and Technology, Syracuse University, 1994. 80p. Paper, \$8 (ISBN 0-937597-38-4).
- Enhancing a New Design for Subject Access to Online Catalogs*. By Karen M. Drabenstott. Ann Arbor, Mich.: School of Information and Library Studies, The University of Michigan, 1994. 162p. Paper, \$12 (no ISBN). Available free via Internet FTP: sils.umich.edu Directory: /pub/papers/ENHANCE. Log on as anonymous with e-mail address or last name as the password.
- Finch, Elsie Freeman. *Advocating Archives: An Introduction to Public Relations for Archivists*. Metuchen, N.J.: Scarecrow, 1994. 202p. \$39.50 (ISBN 0-8108-2935-5).
- Guidelines for Bibliographic Description of Interactive Multimedia*. By the Interactive Multimedia Guidelines Review Task Force. Chicago, Ill.: American Library Assn., 1994. 43p. Paper, \$13.50 (ISBN 0-8389-3445-5).
- Hard Money-Soft Outcomes: Evaluating and Managing the IT Investment*. Concord, Mass.: Paul & Company, 1995. 211p. \$65 (ISBN 1-872474-23-3).
- IBM RISC System/6000: A Business Perspective*. 5th Ed. By Jim Hoskins. New York: Wiley, 1995. 398p. Paper, \$29.95 (ISBN 0-471-04809-7).
- Information Brokers: Case Studies of Successful Ventures*. Ed. Alice Jane Holland Johnson. New York: Haworth, 1995. 117p. \$24.95 (ISBN 1-56024-890-1).
- Information Seeking in Electronic Environments*. By Gary Marchionini. New York: Cambridge, 1995. 224p. \$49.95 (ISBN 0-521-44372-5).
- International Directory of Serials Specialists*. New York: Haworth, 1995. 156p. \$24.95 (ISBN 1-56024-943-9).
- Internet Tools of the Profession: A Guide for Special Librarians*. Ed. Hope Tillman. Washington, D.C.: Special Libraries Association, 1995. 200p. Paper, \$37.50, \$30 members (ISBN 0-87111-430-5).
- A Kaleidoscope of Choices: Reshaping Roles and Opportunities for Serialists*. Ed. Beth Holley and Mary Ann Shible. New York: Haworth, 1995. 386p. \$24.95 (ISBN 1-56024-742-8).
- Mortice Kern Systems. *Internet Anywhere: All You Need To Test Drive Internet Mail and Usenet News for 30 Days*. Disks included. Englewood Cliffs, N.J.: Prentice-Hall, 1995. 430p. Paper, \$29.95 (ISBN 0-13-185612-X).
- Neou, Vivian. *Internet CD*. Englewood Cliffs, N.J.: Prentice-Hall, 1994. 225p. Paper, \$49.95 (ISBN 0-13-123852-3).
- 1994 NFAIS Report Series: Report 1, Changing Roles in Information Distribution*. Ed. Ann Marie Cunningham and Wendy Wicks. Philadelphia: National Federation of Abstracting and Information Services, 1994. 124p. Paper, \$60, \$50 NFAIS members (ISBN 0-942308-44-1).
- 1994 NFAIS Report Series: Report 2, Government Information and Policy: Changing Roles in a New Administration*. Ed. Wendy Wicks. Philadelphia: National Federation of Abstracting and Information Services, 1994. 124p. Paper, \$60, \$50 NFAIS members (ISBN 0-942308-44-1).
- The Online 100: ONLINE Magazine's*

- Field Guide to the 100 Most Important Online Databases.* By Mick O'Leary. Wilton, Conn.: Pemberton, 1995. 256p. Paper, \$22.95 (ISBN 0-910965-14-5).
- Outsourcing Cataloging, Authority Work, and Physical Processing.* Ed. Marie A. Kascus and Dawn Hale. Chicago, Ill.: American Library Assn., 1995. 40p. Paper, \$15, \$13.50 ALA members (ISBN 0-8389-3449-5).
- The Power of Information: Transforming the World: Professional Papers from the 86th Annual Conference of the Special Libraries Association.* June 10-15, 1995, Montreal, Quebec, Canada. Washington, D.C.: Special Libraries Assn., 1995. 94p. Paper, \$25, \$20 members (ISBN 0-87111-442-9).
- Practical Issues in Collection Development and Collection Access.* Ed. Katina Strauch and others. New York: Haworth, 1995. 193p. \$24.95 (ISBN 1-56-24-733-9).
- Preparing a Guide to Your Library and Information Service.* London: Aslib, 1995. 53p. Paper, \$25; \$20 members (no ISBN).
- Proceedings of the Fifteenth National Online Meeting: New York, May 10-12, 1994.* Ed. Martha E. Williams. Medford, N.J.: Learned Information, 1994. 464p. Paper, \$55 (ISBN 0-938734-84-9).
- Proceedings of the 1993 International Conference on Referred Electronic Journals, Winnipeg, Manitoba, Canada, October 1-2, 1993.* Winnipeg, Canada: University of Manitoba Libraries (University of Manitoba, Winnipeg, MB R3T 2N2, Canada), 1994. Paper, \$33, \$32 in Canada (ISBN 0-919932-02-9).
- Public Library Planning: Case Studies for Management.* By Brett Sutton. Westport, Conn.: Greenwood, 1995. 328p. \$65 (ISBN 0-313-28776-7).
- SLA Biennial Salary Survey 1995.* By the Special Libraries Association. Washington, D.C.: Special Libraries Association, 1995. 72p. Paper, \$42 (ISBN 0-87111-423-2).
- Special Librarianship as a Career: SLA Information Kit.* Washington, D.C.: Special Libraries Association, 1995. 100p. Paper, \$20, \$15 members (ISBN 0-87111-40-2).
- Special Libraries and Information Centers: An Introductory Text.* By Ellis Mount. Washington, D.C.: Special Libraries Association, 1995. 272p. Paper, \$37.50, \$30 members (ISBN 0-8711-437-2).
- Silicon Snake Oil: Second Thoughts on the Information Highway.* By Clifford Stoll. New York: Doubleday, 1995. 247p. \$22 (ISBN 0-385-41993-7).
- Stallings, William. *Protect Your Privacy: A Guide for PGP Users.* Englewood Cliffs, N.J.: Prentice-Hall, 1994. 302p. Paper, \$24.95 (ISBN 0-13-0185596-4).
- Teach Yourself Word 6.0 for the Mac* (includes disk). By Stephen Kahn. New York: MIS Press, 1995. 488p. Paper, \$24.95 (ISBN 1-55828-371-4).
- The Value of Corporate Libraries: Findings from a 1995 Survey of Senior Management.* By James M. Matarazzo and Laurence Prusak. Washington, D.C.: Special Libraries Association, 1995. 121p. Paper, \$7.50, \$5 members (ISBN 0-8711-449-6).
- Who's Who in European Research and Development 1995.* New Providence, N.J.: Bowker-Saur, 1995. 800p. Paper, \$315. Offered in combination with *Directory of European Research and Development*. 2 vols. at \$400/set. (ISBN 1-85739-097-0).

“Only Minolta has a copying system that damages books nevermore!”

—EDGAR ALLAN POE

Introducing the Minolta DPCS 3000

If Poe had to copy his original manuscript of “The Raven” the DPCS 3000 would have him waxing poetic. Because it’s the face-up digital publication copying system that copies books gently, simply, quickly.

Thanks to a unique face-up copying design, there’s virtually no spine bending or brittle page breaking. Double-page spreads copy without curves or shadowed gutters. Photos and artwork reproduce to critical acclaim. Even 3-D objects are reproduced with precision.

Digital imaging and laser technology make this a high speed operation. Plus, the DPCS 3000 will soon write a new chapter with image storage, fax, e-mail and Internet service capabilities to meet your expanding needs.

Did we tell you about our service and expertise in copying and digital imaging? We don’t want to give the whole story away. Call 1-800-9-MINOLTA for more information about the most novel idea in copying.

©1995 MINOLTA CORPORATION

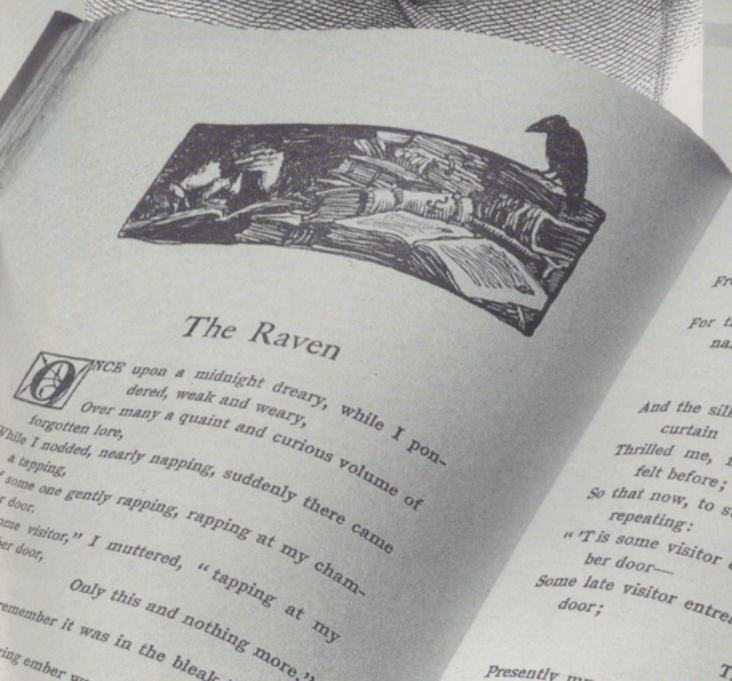
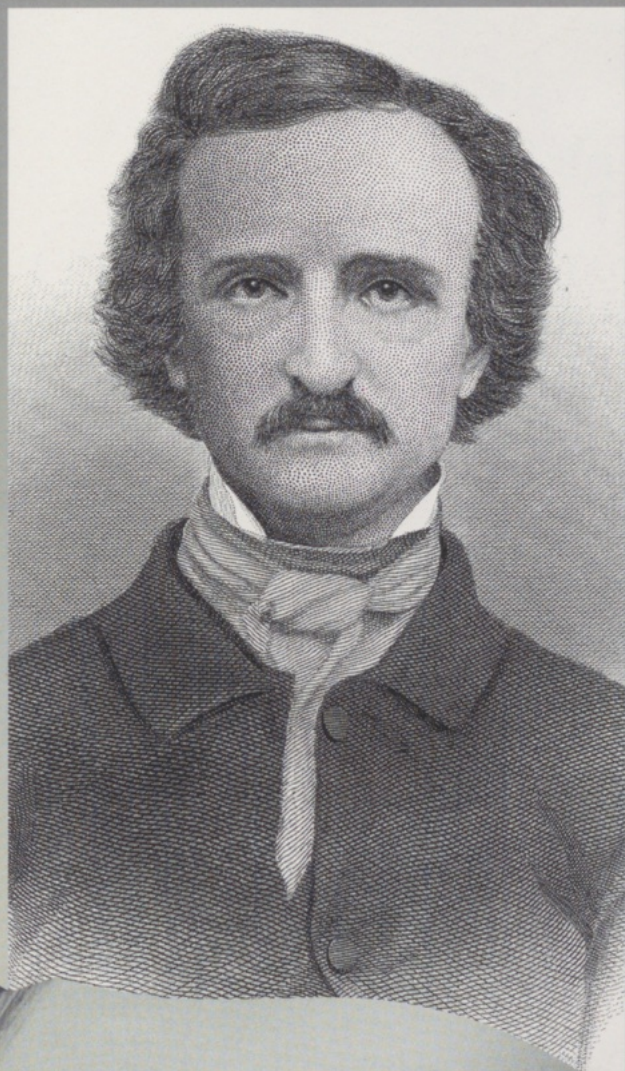
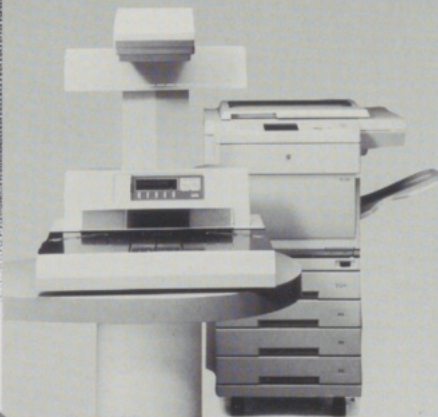
Digital Publication Copying System 3000

Copiers
Cameras
Faxes
Digital Systems
Document Imaging
Camcorders
Binoculars
Color Sensors

Only from the mind
of Minolta



MINOLTA



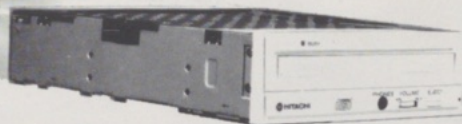
TODD ENTERPRISES INC. NETWORK SOLUTIONS

from 4 to 56+ CD-ROM Drives in ONE CABINET

Todd massive data retrieval Total Access Systems feature dockable double-speed drives and security keylocks. Double speed drives have a three-year warranty. The security lock feature of Todd Total Access Systems is useful in public areas. Todd services include hardware design and development, customized systems, and network installation and management. Todd CD-ROM hardware will operate within your workstation environment with all administrating software. Daisy chain up to 256 CD-ROM drives with the Todd expansion board. Call 800 445-TODD or (718) 343-1040 for more information.



Showing a large 64-drive console and an 8000 series tower. Consoles available in five different sizes. Both consoles and tower with internal CPU units and dockable drives. Network towers, multidrives and consoles with your existing workstations.



Todd dockable drives slip in and out without tools or wiring. It is easy to expand your system.

since
1971
over 20 years
OF EXCELLENCE
TODD
Enterprises

224-49 67th Avenue · Bayside, NY 11364
718 343-1040 · FAX 718 343-9180 · BBS 516 829-0212

800 445-TODD

available on GSA contract