

# Data-Guided Collection Development: A Promise Unfulfilled

Dennis P. Carrigan

The widening gap between constituents' demands for materials and research libraries' abilities to meet those demands calls for new strategies or emphases by those libraries. It may be expected that patron demands will exert increasing influence over the allocation of materials budgets and those data on use of materials will drive or guide collection development. Are such data, which can be provided by automated systems, being used to guide collection development? Responses to a survey reveal that, with few exceptions, they are not.



Among libraries serving the leading North American research universities, there is clear evidence of a widening gap between constituents' demands for materials and libraries' abilities to meet those demands from their own collections. According to the report on the *ARL/RLG Interlibrary Loan Cost Study*, interlibrary borrowing among ARL libraries grew by 108 percent in the ten years through 1992, or at an average annual compound rate of 7.6 percent. The increase the final year was ten percent.<sup>1</sup> Moreover, a graph in the 1993-94 *ARL Statistics* portrays vividly the near-doubling of interlibrary borrowing among ARL academic libraries between 1986 and 1994 (see figure 1).<sup>2</sup>

According to the cost study, three converging trends explain the increase in interlibrary loans. These trends account for the widening gap between constituents' demands and libraries' abilities to meet those demands. They are "more accessible and easy-to-use bibliographic tools; . . . a growing universe of published items";

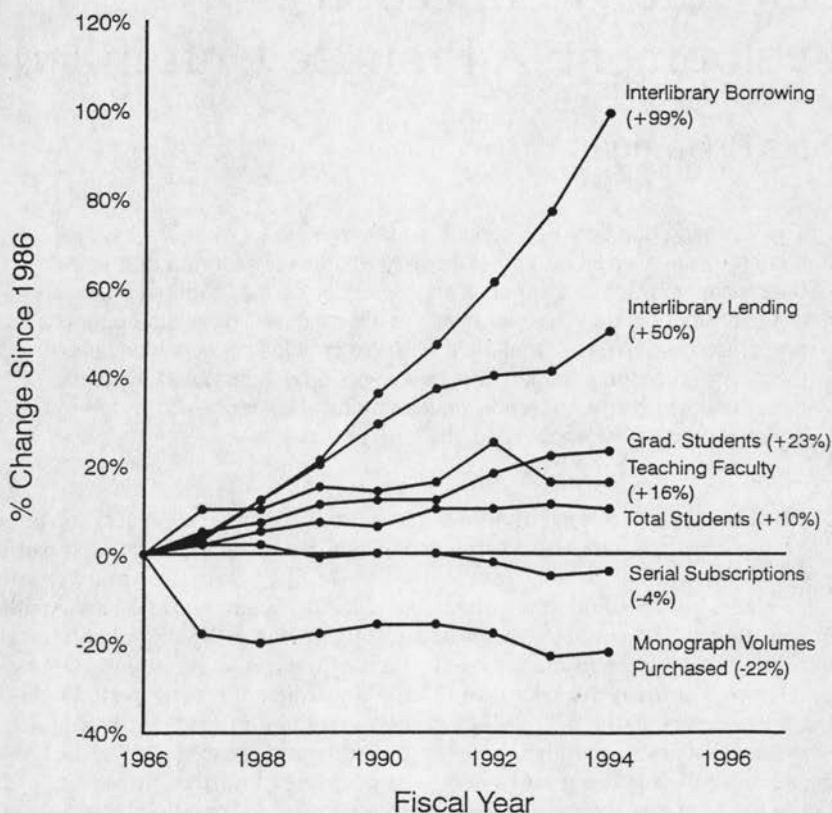
and "reduced buying power" for libraries as a result of increased acquisition costs for most research resources combined with constrained budgets.<sup>3</sup> Another graph from the *ARL Statistics* portrays the third of the converging trends. It reveals the disparities, during the period 1986-94, between the unit price increase for serials and monographs (115% and 55%, respectively), and the slower rate of increase in library spending for serials and, in particular, for monographs (93% and 17%, respectively). The graph also shows the consequences of these disparities, a 4 percent decline in serials purchased and a 22 percent decline in monographs purchased (see figure 2).

The graphs depict what Nancy L. Eaton refers to as the "crisis of scholarly publishing," which has had severe consequences for research university library budgets.<sup>4</sup> As the director of such a library, Eaton asserts:

Pleading for more money to match inflation and the fluctuations in the

Dennis P. Carrigan is Assistant Director, School of Library and Information Science, University of Kentucky; e-mail: carrigan@ukcc.uky.edu.

**FIGURE 1**  
**Supply and Demand in ARL Libraries, 1986-94**



Note: This graph compares growth in numbers of users with changes in onsite resources and resource-sharing.

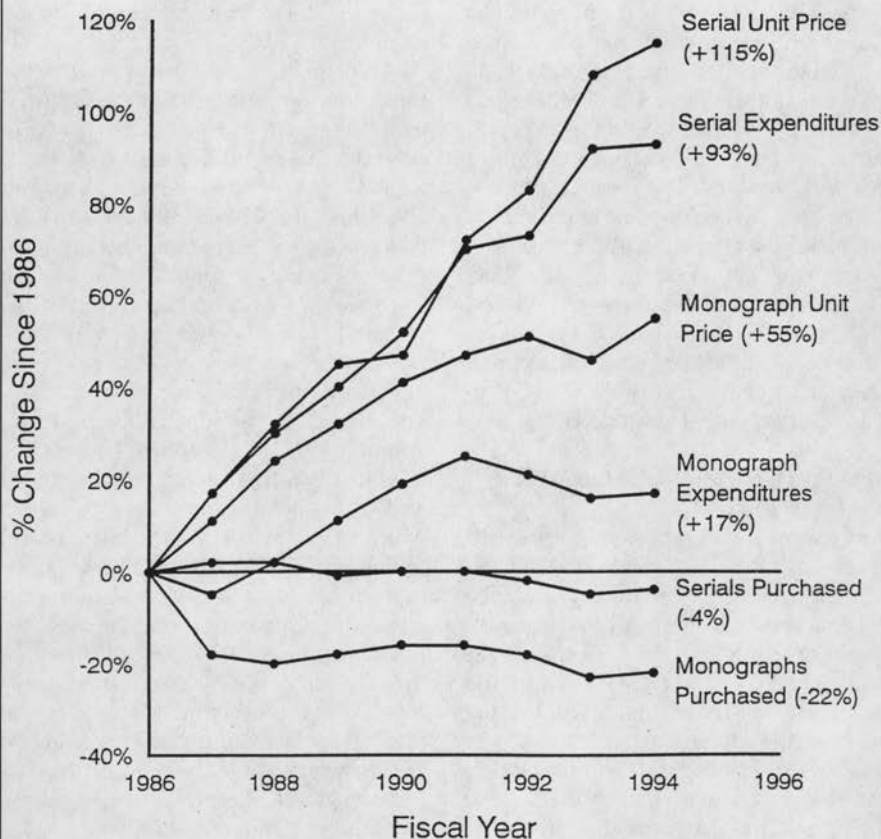
Source: 1993-94 *ARL Statistics*. Prepared by Kendon Stubbs; copyright © 1995 by the Association of Research Libraries.

dollar is no longer viable unless it is also accompanied by new strategies that hold promise for changing the system. Library directors must also question how well we are using the resources we have. The old 80/20 rule of thumb—20 percent of our collections meet 80 percent of demand—is no longer credible to our senior [university] administrators and funding agencies, given the fiscal climate.<sup>5</sup>

#### **Importance of Use Data: Increased Significance of Demand**

Charles B. Osburn believes libraries are establishing a new set of guiding principles in response to persistent and significant stresses and strains, ambiguities and change, which are either triggered or aggravated by economic pressures. One of the principles is a shift in emphasis from spending acquisitions budgets on the basis of speculation to spending on the basis of demand.<sup>6</sup> Establishment of such a

**FIGURE 2**  
**Monograph and Serial Costs in ARL Libraries, 1986-94**



Source: 1993-94 ARL Statistics. Prepared by Kendon Stubbs; copyright © 1995 by the Association of Research Libraries.

principle should increase the importance of the data that can be produced by automated systems. These data reveal the demand for materials and could guide collection development and thus the spending of acquisitions budgets.

### Questionnaire

To determine the extent to which data that can be produced by automated systems are used to guide collection development, the author developed a questionnaire and sent it to the chief collection development officer at each of the 108 university library

members of ARL. Respondents returned seventy-nine completed questionnaires, for a response rate of 73.1 percent. Respondents were asked if they could be quoted or paraphrased by name; not all consented.

### Discussion of Responses

#### *Universality of Automated Systems: Use of Data Leads to Better Decisions*

The first question asked respondents if the main library uses an automated circulation system. Not surprisingly, all respondents replied affirmatively. Respon-

dents were then asked if the data produced by the system are regularly used in collection development decisions. Thirty-four (43%) replied affirmatively. Those who replied affirmatively were asked if they believe use of the data leads to significantly better collection development decisions than would be the case if the data were not used. Twenty-eight of the thirty-four (82.4%) replied affirmatively. Those who reported that the data produced by an automated circulation system were *not* regularly used in collection development decisions were asked if it is planned to use such data for such purpose. No time horizon was specified. Thirty-two of the forty-five (71.1%) replied affirmatively (see table 1).

#### *How the Data Are Used: Macro-Level Use Infrequent*

Respondents who reported that their libraries use the data in collection development decisions were asked to describe the data and how they are used. Their responses fell into two categories. One category has to do with use of the data at the micro level, or individual title level; the other at the global (macro), or subject classification (call-number range) level.<sup>7</sup> Only eleven respondents reported use of the data at the macro level. That is 32.4 percent of those who reported their libraries use the data in collection development decisions, but only 13.9 percent of all

respondents. However, use must be at the macro level if the data are to bear significantly on resource allocation and guide collection development<sup>8</sup> (see table 2 for a summary).

At the micro level, twenty-four of the thirty-four respondents (70.6%) reported using information about title holds to consider purchasing additional copies. Fourteen (41.2%) reported using information about lost or damaged titles to consider purchasing replacements. Eight (23.5%) reported using circulation data to make storage decisions and eight institutions use such data to determine candidates for withdrawal. Five (14.7%) said data are used in canceling journal subscriptions. At the macro level, six respondents (17.6%) reported using the data in allocating the materials budget and five said data are used to determine circulation by classification, which information may influence budget allocation. As approval plans increase in popularity, it may be especially interesting to other institutions that three respondents (8.8%) reported use of the data to monitor and adjust approval plans. Because of comments made in two of the approval plan responses, the author decided to treat the three as macro-level use of the data.

Several respondents commented on the difficulty that a NOTIS system presents. According to Chuck Hamaker, assistant dean for collection development at Louisiana State University, "One of the major problems for most libraries with NOTIS is that it takes special programming to get the SASS reports that coordinate data to make it meaningful." Another respondent at a NOTIS library commented that purchase of NOTIS Quick Reports enabled the library to produce some reports. However, Quick Reports "was very expensive

**TABLE 1**  
**Summary of Responses to Questionnaire**

1. Questionnaires sent	108 (100.0%)
2. Questionnaires returned	79 (73.1%)
3. Automated circulation system	79 (100.0%)
4. Data regularly used in collection development decisions	34 (43.0%)
5. Use of data leads to significantly better decisions	28 (82.4%)
6. Data <i>not</i> regularly used in collection development decisions	45 (57.0%)
7. Plan to use data in collection development decisions	32 (71.1%)

**TABLE 2**  
**How the Data Are Used**

Level of use	Micro	Macro
1. Add copy	24 (70.6%)	
2. Replace copy	14 (41.2%)	
3. Store	8 (23.5%)	
4. Withdraw	8 (23.5%)	
5. Cancel journals	5 (14.7%)	
6. Allocate budget		6 (17.6%)
7. Determine use by classification or range		5 (14.7%)
8. Monitor approval plan		3 (8.8%)

ability of those specific titles most in demand by current students and faculty. A relatively small amount of money spent here can result, I believe, in a disproportionate benefit to customers. . . . Also, as we face another round of journal cancellations, use data for sci/tech journals captured at relatively low cost will aid greatly in the decision process.

[and] there are some very annoying omissions."

Conversely, Linda Di Biase, collection development librarian at the University of Washington Libraries, which uses an Innovative Interfaces Integrated System, noted the system's excellent management reporting capabilities.

#### *"Significantly" Better Decisions*

Does the use of data lead to significantly better collection development decisions than would be the case if the data were not used? It does, according to twenty-eight of the thirty-four respondents (82.4%) who reported the data are used. Those who answered yes were asked to elaborate, though not all did so. Jim Kuhlman, associate dean of libraries for collections and information services at the University of Alabama, commented on use of the data at the micro level:

"Significantly" is a relative term. We haven't made the use of circulation data that I would have hoped. For example, it has not played a role in budget allocation other than helping to make the case for increasing the amount for replacement/added copies. On the other hand, I think that has been significant. Such data has provided [an] accurate and convenient mechanism to increase the avail-

At Auburn University, Yvonne Kozlowski expressed the view of several respondents who use the data at the macro level: "We know what areas of our collections are receiving high use." As a result, more materials are acquired for those areas. At the University of Illinois-Urbana, Karen Schmidt, chair of the Collection Development Committee, captured the essence of what use of the data makes possible: "It allows us to make informed decisions about where our dollars should go."

Suzanne Pitman, associate director for technical services and collection management at North Carolina State University, made a number of good points that should be of interest to her colleagues, especially those not using data generated by an automated system:

The use of hard data, as opposed to anecdotal or individual impression, leads to more reliable analyses; data is now available that wasn't in the manual system, and data are routinely available now that formerly required special efforts. The availability of data over time helps analyze trends more accurately. The ability to tailor reports supports more extensive and creative use of data to better understand use patterns and our user communities.

### *Explaining Lack of Use*

The forty-five respondents who reported that their libraries use an automated circulation system, but that data produced by the system are *not* regularly used in collection development decisions, were asked to explain the lack of use. Not all respondents did so; several respondents cited more than one reason (see table 3).

System limitations were cited by eleven of the forty-five respondents (24.4%). At Arizona State University, Elliot Palais, coordinator for the social sciences, stated "that a number of years ago the library adopted the CARL automated system, which provides only the most rudimentary kind of data on collection use. The library hopes to install a new system within a year or so, and a requirement for it will be the ability to produce a variety of collection use statistics."

Barbara Van Deventer, assistant director for collection development at the University of Chicago, commented that the system in use there is a local system developed more than a dozen years ago. Although certain categories of information useful to patrons are provided, nevertheless "data retention helpful to collection management was not included [in system design]. Thus, we have no information on use through circulation. This is a serious defect."

According to one respondent, "Unfortunately, the libraries' present circulation system does not make data that is useful for collection development readily available." And another respondent noted, "Our NOTIS system yields gross circulation data presently. We are

not receiving the Quick Reports which would provide useful collection management data." A third respondent added: "The data that will be generated [from a new automated circulation system] will not be sufficiently attuned to collection management needs for it to be helpful."

At the University of Kentucky, Bonnie Jean Cox, collection development librarian, reported, "The system's data cannot be accessed from the desktop. Producing such a report requires the intervention of a systems librarian and a programmer." At MIT, according to Carol Fleishauer, associate director for collections services, "It's too difficult to extract the data from the systems." Lynn Sipe, acting director of libraries at the University of Southern California, wrote that circulation data now are used for replacement and added copy decisions, and "more sophisticated use would be made *if* the automated data were more easily retrieved, especially without assistance of computer room staff."

Another eleven respondents reported having data for too few years. According to Joel Rutstein at Colorado State University, his institution has had three systems since 1985, and, as a result, staff have been unable to accumulate reliable information over sufficient time. The current system has been in place only two years. Another respondent wrote that her institution has "only a few years of data that indicate usage, etc., in our present SIRSI circulation system."

Seven of the forty-five respondents (15.6%) reported that within the context of automated systems, other programming had higher priority. At Cornell University, according to Ross Atkinson, associate university librarian, the necessary programs have not been written because they have been "a lower priority than other programs needed for processing and public access." The head of collection development at a Canadian institution also commented that special programming is needed and "our library system has given

**TABLE 3**  
**Why the Data Are Not Used**

1. Limitation of system	11 (24.4%)
2. Data for too few years	11 (24.4%)
3. Needed program(s) unavailable	7 (15.6%)
4. Not convinced of value of data	9 (20.0%)
5. Other	7 (15.6%)

little priority to deriving management information." At the University of Hawaii at Manoa, according to Carol Schaafsma, collection development coordinator, "We have not had the necessary expertise to extract the information being collected. We are in the process of hiring an individual who will be charged with responsibility for doing this."

Joseph Branin, associate university librarian at the University of Minnesota, commented, "We automated circulation late (about five years ago) and other automation issues have taken priority over programming for useful statistical reports."

At some institutions, however, the reason for not using the data rests with collection development itself. At one of the leading U.S. public institutions, according to the respondent, the "basic problem is that collection development has not conceptualized and requested data that might be helpful." Another respondent wrote that there is "no particular reason" why the data are not regularly used in collection development decisions. And at a major Midwestern institution, it was reported that "It's something we haven't gotten to on a regular basis yet."

At Washington University in Saint Louis, according to B. J. Johnson, head of collection development services, consideration has been given to using circulation data in collection decisions, "but effort to produce reports from the circulation system is much more trouble than information is worth." Johnson then makes an interesting point: "Also, local history and politics would not tolerate any significant changes—regardless of source of data!"

Several respondents asserted that circulation data are not seen to be important for guiding collection development due to the nature of their institutions. At one Midwestern institution, the head of collection development commented that "because of the research aspect of [our] collection and for political reasons circ data is a 'hard sell.'" The collection development officer at another Midwestern

institution commented, too, on the influence exerted by the institution's research mission: "Circulation isn't viewed as the most important factor in collection development in a research library."

The head of the Collection Development Department at a Canadian institution also noted the influence exerted by the need to support research: "Our research collection is a single-copy collection. This means that the number of times an individual title circulates over a given

---

**Several respondents asserted that circulation data are not seen to be important for guiding collection development due to the nature of their institutions.**

---

period is largely irrelevant." Another respondent made a similar observation, saying that circulation "is a minor criterion for evaluating collections" at a research library and "takes no account of in-house use, ILL, or use of special collections." Yet, circulation could be defined so as to take account of in-house use and ILL.

Finally, the comments made by two respondents may contain valuable warnings. According to one, "The effectiveness of the data is related only to the actual use of it and that is entirely dependent on the ability and the willingness of the selectors to take advantage of what can be made available." In that regard, Jim Kuhlman at the University of Alabama, whose library *does* use the data in collection development decisions, expressed the view that "many librarians are uncomfortable with statistics, especially if the resulting conclusion contradicts our preferences."

### **Concluding Comments**

This article reports findings concerning the use of the data available from automated circulation systems to guide collection development. Although it is be-

lieved that such data are valuable in guiding collection development, the data are not the only information that is valuable to collection developers. The importance of interlibrary loan statistics in collection development has been asserted. Moreover, those who contend a formula is the best means to allocate the materials budget argue, if only implicitly, that a considerable amount and variety of information should bear on collection development. In one formula, for example, information was incorporated on undergraduate enrollment, number of graduate majors, number of faculty by department, holdings by subject, and average price by subject.

In spite of the value of other data to collection developers, however, some have predicted that it is the data available from automated circulation systems that would guide collection development. Fifteen years ago, H. William Axford wrote that there was "every reason to believe that the long-term collection use study will become widespread and that it will result in fundamental changes in the way library collections are managed

---

**Some respondents lamented the nonavailability of needed programs and the inability to persuade the appropriate persons that the programs should be developed.**

---

in the decade ahead." Economic constraints would be the principal impetus for the use studies. He continues, "This unpleasant [economic] reality poses an impossible budgeting problem for institutions of higher education unless there is a willingness to explore new approaches to resource allocation which incorporate the concept of attempting to measure the return on investment." Senior campus administrators and external agencies would see that the institutions' libraries were not exempt from the requirement to explore new approaches to resource allocation. As one result, collection use stud-

ies, made possible by "sophisticated automated circulation systems, would become common"; and according to Axford, "it is highly probable" such studies "will prove to be the most effective tool for . . . upgrading the management of academic libraries to emerge from a decade of experimentation."<sup>9</sup>

Nearly a decade ago, Susan Nutter wrote:

Collection development librarians are facing increasing pressures to tailor collections more precisely to user needs and to do so in a more cost-effective manner. Increasingly, collection management decisions need to be made on the basis of hard data—assumptions, educated guesses, and intuition are no longer adequate.

Fortunately, Nutter continued, "with this growing need, computer-based systems that can provide this management information are coming into widespread use in libraries."<sup>10</sup>

Charles Hamaker wrote of "fiscal reality" and "accountability," and of the potential benefits available from computer-based systems: "As many libraries, large and small, begin to utilize data from online circulation and control systems . . . one of the least examined areas has been the potential impact on collection development and management decisions." According to Hamaker, other demands stand in the way of having the potential benefits realized, but he believed it was inevitable that "collection design" would come about and would lead to the benefits, which are a better match between collections and user needs. The selection "process is about to change even in some of the largest libraries. Fiscal reality will drive the change and accountability will become a necessity. . . . With the availability of circulation information with just a few keystrokes, the question of which book to add has a whole dimension available that was guesswork before."<sup>11</sup>



Yet, more recently, Hamaker has written that although "OPACs held [and hold?] the promise of better management data, more information on our collections, [f]or the most part this information has not been collected, analyzed, utilized. It is a promise yet to be fulfilled."<sup>12</sup>

The results of this survey bear out that assessment. The predictions of Axford and Nutter have not been borne out. Why, in spite of their predictions, is so little use made of data that can be provided by automated circulation systems to guide resource allocation and collection development? The responses to the survey reveal a variety of reasons. System limitations were cited by a number of respondents and seemed to be a particular problem at NOTIS sites. Some respondents lamented the nonavailability of needed programs and the inability to persuade the appropriate persons that the programs should be developed. Still other respondents made it clear they are not convinced of the value, for collection development, of data available from automated circulation systems.

Is there, however, another, overarching explanation, not revealed in returned questionnaires, for the nearly universal nonuse of the data at the important macro level? In discussing, and in some cases predicting, use of the data "to shift resources internally" (Eaton), to bring about "fundamental changes in the way library collections are managed" (Axford), or "to tailor collections more precisely to user needs" (Nutter), reference was made to an expected increased concern on the part of university administrators and funding agencies about how libraries allocate their resources. Perhaps that has not come about. Perhaps, for the most part, libraries serving research universities have not been subjected to the increased accountability that some saw coming. If the anticipated heightened accountability has not materialized, then a powerful incentive for a shift in emphasis "from spending acquisitions budgets on the basis of speculation to spending on the basis of demand" (Osburn) has not developed. This issue is of such significance that it deserves further exploration.

---

## NOTES

1. Marilyn M. Roche, *ARL/RLG Interlibrary Loan Cost Study* (Washington, D.C.: ARL, 1993), 1.
2. Martha Kyrillidou, Kaylyn E. Hipps, and Kendon Stubbs, eds. & comps., *ARL Statistics, 1993-94* (Washington, D.C.: ARL, 1995), 8.
3. Roche, *ARL/RLG Interlibrary Loan*, 1.
4. Nancy L. Eaton, "The Director's Role in the Acquisitions Dilemma," *Journal of Library Administration* 19, no. 2 (1993): 3.
5. *Ibid.*, 4-5.
6. Charles B. Osburn, "Collection Evaluation and Acquisitions Budgets: A Kaleidoscope in the Making," *Journal of Library Administration* 17, no. 2 (1992): 5.
7. I am indebted to Paul Metz, Virginia Polytechnic Institute and State University, for the terms *micro* and *global*, which he used in his response.
8. Data presented in a table may seem to be inconsistent with data reported in the text due to multiple responses of a single respondent.
9. H. William Axford, "Collection Management: A New Dimension," *Journal of Academic Librarianship* 6 (Jan. 1981): 325-26, 328-29.
10. Susan K. Nutter, "Online Systems and the Management of Collections: Use and Implications," in *Advances in Library Automation and Networking*, ed. Joe A. Hewitt (Greenwich, Conn.: JAI Pr., 1987), 126.
11. Charles A. Hamaker, "Management Data for Selection Decisions in Building Library Collections," *Journal of Library Administration* 17, no. 2 (1992): 71.
12. Charles A. Hamaker, "Time Series Data for Collection Development or: You Can't Intuit That," *Library Acquisitions: Practice & Theory* 19 (summer 1995): 195.



# **FirstSearch<sup>®</sup>**

*A world of information online*

## THE FASTEST GROWING END-USER DATABASE SERVICE IN THE LIBRARY COMMUNITY

- Full-text *and* images\* Online
- Web Access
- NetFirst<sup>™</sup>, the authoritative guide to Internet resources
- WorldCat<sup>™</sup>, the end-user version of the OCLC Online Union Catalog
- More than 50 databases

**To find out how FirstSearch<sup>®</sup> can help you move  
to the next stage of the electronic library:**

Call us at 1-800-848-5878, ext. 6425 to request a copy of  
the new FirstSearch brochure. Or view and order it online through our  
home page on the World Wide Web:

*<http://www.oclc.org>*

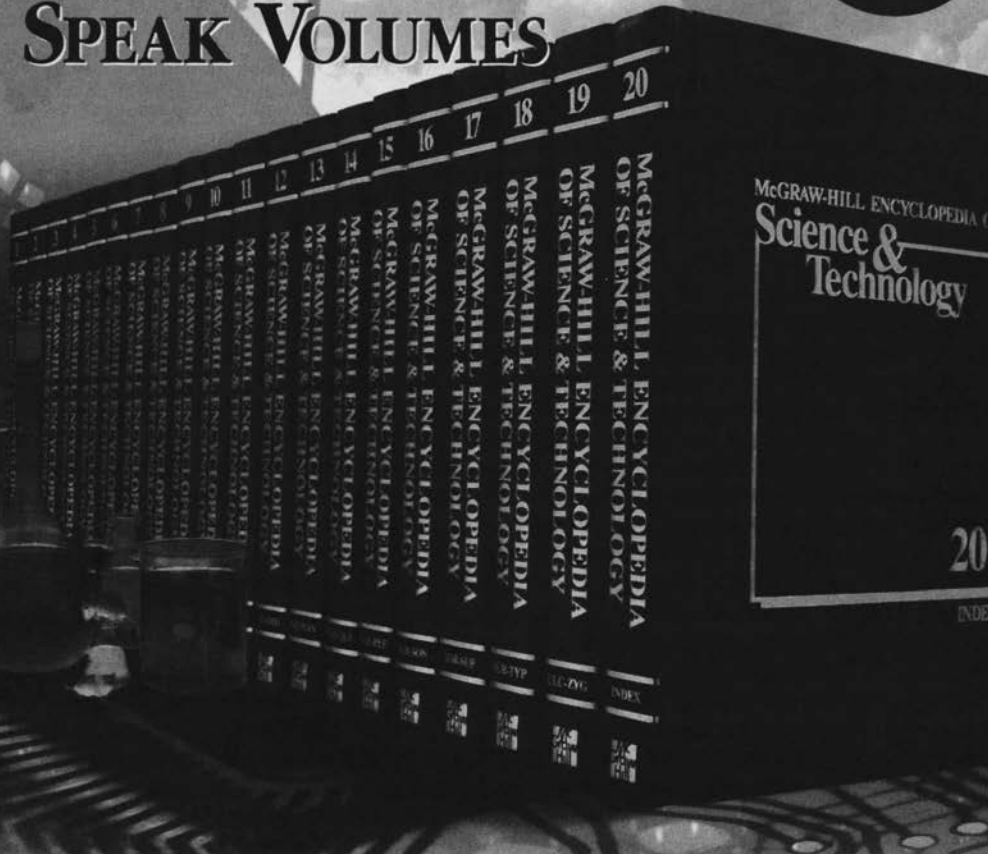
\*Coming soon



FURTHERING ACCESS TO THE WORLD'S INFORMATION

Coming  
May 1997

# SOME SCIENTIFIC ACHIEVEMENTS SPEAK VOLUMES



## Announcing McGraw-Hill Encyclopedia of Science & Technology, 8th Edition

The sheer volume is staggering. Seven thousand five hundred science and technology articles from renowned contributors, including 19 Nobel laureates ... 1,600 new articles ... 13,000 photographs ... 20 volumes. Being so authoritative, accurate, reliable, and comprehensive is a tall order. But we can handle it. You wouldn't expect anything less from the ultimate science and technology reference.

**Special Prepublication Offer—  
Save Over \$200**  
Pre-publication price \$1,750 through 5/31/97  
\$1,995 thereafter. ISBN: 0-07-911504-7

Available through all library wholesalers or call 800-722-4726; fax 614-755-5645  
For a brochure, call Rebecca Seger at 212-337-5036; e-mail [rseger@mcgraw-hill.com](mailto:rseger@mcgraw-hill.com)  
[www.books.mcgraw-hill.com](http://www.books.mcgraw-hill.com)

A Division of The McGraw-Hill Companies

