

A Study of Collection Overlap in the Southwest Indiana Cluster of SULAN

**Ruth H. Miller, Collection Development
and Preservation Librarian
Cunningham Memorial Library, Indiana State University
Terre Haute, Indiana**

**Martha W. Niemeier
Collection Development Librarian
University of Southern Indiana, Evansville, Indiana**

Indiana's State University Library Automation Network—SULAN—at this stage is most fully developed in the Indiana State University (Terre Haute) "cluster." Rose-Hulman Institute of Technology and St. Mary-of-the-Woods College, also in Terre Haute, have been part of ISU's NOTIS system since 1985/86; the University of Southern Indiana at Evansville and Vincennes University, Vincennes, are scheduled to be added to this cluster in 1989/90. Corresponding to the need for this statewide system (eventually to link the major academic libraries in the state) is a need to study collections and plan for meeting future needs. We must know as fully as possible what we have before appropriate and consistent decisions can be made on resource sharing.

No one now questions the need for resource sharing and cooperation among libraries; it is an essential fact of life. But as libraries join networks and share automated systems, how much sharing and what kind of cooperation have to be worked out more precisely, balancing local autonomy with meaningful cooperation. Some of the deterrents

to cooperation may be differing library sizes and types, the amount and kind of technology available, the history of each library and school, distance between libraries, state and local politics, and the focus of campus or library administrations. Librarians working in networks need to be sensitive to such potential problems and work to prevent or resolve them.

As Ionesco notes, "cooperation in collection development is going to be conducted on a regional, rather than on a national basis," which "should result in more efficient use of common resources." In addition to political and psychological obstacles to cooperation, she mentions such objective impediments as lack of knowledge of the collections.¹ Many major research libraries have used the RLG Conspectus to address this need, and from the knowledge gained have shaped comparative data critical to making hard choices among specialties. Though it may be helpful as a way of assessing relative strengths and comparing relative holdings, for the medium- and small-size library the Conspectus provides limited assistance. As Nisonger points out, "The

level of specificity in the RLG Con-spectus breakdown may be too detailed for libraries whose collec-tions are smaller than the size of ARL."² Others have discussed cooperative collection development among Indiana's three largest librar-ies—Indiana University, Purdue, and Notre Dame³—but collections of the state's smaller libraries need to be analyzed also because they, too, have much to contribute.

While it is true that libraries "are moving away from their past emphasis on collection building and growth to a new emphasis on pro-viding access to information from many sources," there is a continuing need for strong core collections to serve the local constituency and curriculum. "Resource sharing will be more necessary in large reserch universities, where the demand for little-used and more esoteric materi-als will be greater. Small and medium-sized libraries will still need to devote the bulk of their resources to collection development" because that will be required to provide for undergraduate needs.⁴ Small and medium-sized libraries continue to have a different set of needs for collection development than major research libraries. Until a collection reaches a rather extensive size and complexity, access to off-campus materials doesn't help undergradu-ates much, because their needs are more basic and immediate. Such access is of much greater use to faculty and graduate students who may work over longer periods of time, for whom a specific citation or type of information (rather than any of several sources) is crucial, and who already have the basic tools of research available.

Balancing the needs to provide materials for the undergraduate curriculum and to assist more advanced researchers calls for what

Mosher describes as "collaborative" collection development, which "requires communities of librarians who know their patrons and their programs well, and who are active rather than passive in their ap-proach to collections." The Conoco Study "surveyed groups of bibliogra-phi-ers from the humanities (German literature) and sciences (Geology) to examine the degree to which coop-erative collection development could be effective" German literature bibliographers were willing to change 40%, and geology bibliogra-phi-ers up to 50% of their selection decisions and rely on collections elsewhere "if they could be reason-ably sure of both bibliographic access and physical availability of items in those collections"⁵

As knowledge of what is avail-able becomes more widespread and access to it more rapid, that a particular title be actually available in a collection at any given time becomes less critical. So long as some library has the title and it can be shared quickly, it need not matter who owns the title provided undergraduate needs are being met with appropriate alternative sources. Selection based on journal indexes, core lists and bibliographies may become less important just because the titles included are among the most likely to be purchased by other libraries and thus to be readily available elsewhere. Medium and small-sized libraries will do what they have always done but with even more assurance—select from the core lists only those few items most needed to support the curriculum and most likely to be used by local patrons. If a title is central to the curriculum, it should be available locally when needed for reserve or multiple checkout. And since each library must proceed with orders as best it can, obviously there will be

some unintentional overlap because no library can wait for months to see if others will acquire each title. The greater difficulty is determining the boundary between what is necessary for the curriculum and what are more specialized needs, as well as determining who will attempt to cover the latter at what level of comprehensiveness. Attempting to acquire even the "necessary" volumes probably will continue to be beyond our financial means.

What difference does collection overlap make? Though the most efficient use of state funds might appear to be having only a single copy of each title in the state, in fact that would ensure an overwhelming burden on Interlibrary Loan and force many patrons to wait too long for materials. Some overlap is desirable, especially with high use items, because libraries can afford few multiple copies. In addition to occasional high use, there is also some need for overlap to compensate for theft and mutilation, particularly considering how many of those titles may be irreplaceable because no longer in print. So a certain amount of redundancy is needed within each collection and certainly within a region. Yet we also want the greatest possible variety, range, and depth to the collections in order to gain the interdisciplinary strength that few individual collections have alone.

Method/Rationale

To develop a base from which to begin our study of any possible cooperative efforts, we took a collection sample, focusing primarily on Indiana State University, Terre Haute, and the University of Southern Indiana, Evansville. Since the holdings of Rose-Hulman and St. Mary's are available online with

ISU's holdings, their data is included where readily available in order to give a fuller picture. Once USI's and Vincennes University's holdings are added, future studies may include data on all five libraries.

Several overlap studies⁶ have been done using OCLC tapes. While this method may be more complete, it is also more costly than this preliminary study required. We wanted a sample check to give some idea of what to expect as the network becomes a reality. The goal of the project is knowledge of quantitative collection overlap and uniqueness in order to develop qualitative criteria for future use.

There is no intention at this stage to change collection development policies of either library, though some adjustments may eventually be possible in conjunction with fuller statewide cooperation. The kind of delivery system, for example, that is worked out—whether UPS, U.S. mail, bus, cooperative van, or some other means—will affect availability, cost, effectiveness, and convenience. Those factors could, in turn, affect many specific collection development decisions.

We used *Books for College Libraries*, 3rd ed. because of its status as a basic collection. Each school had an equal opportunity to be aware of the titles and to choose them for purchase. The limitations of funds and differing curriculum emphases account for many of the differences in the choices. Not all titles even from a core list are appropriate for all libraries, but these titles were among the most likely to be chosen by these libraries. Significant overlap was expected. Academic libraries of well over one million monographic

volumes would very likely have almost all of these titles. ISU has a collection approximately five times the size of USI's collection. Aside from the difference in size between ISU and USI, it should be noted that USI (as ISUE) was until 1985 a sister campus to ISU, though the libraries never coordinated collection development. The percentage of volumes held would have increased if we had counted every edition held; instead we counted only the edition listed in BCL3 or a reprint of that same edition. A random sample of 497 titles was selected from BCL3,⁷ and those titles were checked to determine holdings and "gaps," uniqueness and overlap in the holdings of ISU, USI, Rose-Hulman, and St. Mary's. Based on Shaw's finding that both publication date and subject were predictors of overlap, we look at these factors also.⁸

A second part of the study was selecting seven LC sections, samples roughly corresponding to LC classes represented in the five volumes of BCL3, and checking the complete shelflist holdings of USI in these sections against the online holdings of ISU. This shelflist count was not meant to be considered any kind of core list or bibliography of notable books like the BCL3 count. Instead it was intended to give a different kind of sampling, showing the depth of collection overlap in several small areas. This was intended to suggest whether the relative percentages derived from the BCL3 sample are comparable to the rest of the libraries' collections. In both cases we wanted to see how much duplication there is in numbers and in percentages and where the duplication is to be found.⁹ By taking the BCL3 sample and the shelflist sample we were able to use two of the four sampling methods Buckland *et al.*

list: comparison of segments of the catalog and sampling from external lists.¹⁰

Findings

Of the 497 items checked in BCL3, 105 (21.1%) were held by none of the four libraries. The remaining 392 (78.9%) were held by at least one library. Of those, 222 (44.7%) were held by only a single library: 192 (38.6%) by Indiana State University; fifteen (3%) by the University of Southern Indiana; ten (2%) by Rose-Hulman; and five (1%) by St. Mary-of-the Woods. Of the remaining 170 items (34.2% of the sample that overlap), 128 (25.8%) were held by two libraries; 32 (6.4%) were held by three libraries; ten (2%) were held by all four libraries. (ISU had 72.8% of the sample titles; USI had 29.2%; St. Mary's, 12.1%; Rose-Hulman, 9.7%.) Nisonger found 51.9% of titles in his study of seventeen Texas libraries to be held by a single library, but he states that in this count "uniqueness was overestimated and duplication underestimated."¹¹ Our finding on a more limited study was that 44.7% of titles were unique. (See Chart 1.)

Of the 105 titles not available from ISU, USI, Rose-Hulman, or St. Mary's, an OCLC check indicated that 52 (10.5% of the total BCL3 sample and 49.5% of the "gap" titles) are available from two or more libraries that will eventually be part of SULAN. Another 27 (5.4% of the total sample and 25.7% of the gap) are available from at least one library that will be part of SULAN. Nine (1.8%, total and 8.6%, gap) others are available within the state but not from a "network" library, and 17 (3.4%, total and 16.2%, gap) are apparently not available within the state. (See Chart 2.) If only 5.23% have to be requested from

non-network libraries, then apparently the network could satisfactorily provide for patron requests except those unable to wait even a few days.

We did not restrict this sample by date, but we did look at dates of publication for the BCL3 sample. Shaw's study of public and academic libraries in Indiana found that the "highest overlap occurs among books published in the 1960s."¹² Our survey showed that for BCL titles our actual number of volumes that overlap was greater in the 1970s (59 of 180 or 32.8%), but the percentage was higher for the 1960s (48 of 99 or 48.5%). This is interesting since USI has only existed since 1965, but their funding—like that of most other academic libraries—was strong from 1965 until the early 1970s, and during that period they would have emphasized 1960s publications in their purchases. For the 1950s 20 of the 50 titles overlap (40%), and for the 1980s 30 of 137 overlap (22%). Chart 3 shows the BCL3 sample breakdown into decades and the numbers that overlap in each.

Shaw also found that the "books from academic libraries were primarily in the social sciences (46.4% of the academic library sample)."¹³ Our survey showed that we have a higher overlap in the arts and humanities, followed by overlap in the social sciences, if we group the LC classes as Shaw does. When we group by the categories used in BCL3, the LC class/BCL volume with the highest overlap is P (literature) with 44.7%, followed by the G-L classes/volume, with 33.6%, and the third highest category is the C-F classes/volume with 30.1% overlap. Chart 4 shows the percentage of the sample that overlaps. LC classes A, C, S, and V showed zero overlap. LC class U shows 100% overlap, but

that means less when it is noted that this sample includes only two titles. Our study is more limited than Shaw's, and a comparison of our total holdings could, of course, produce somewhat different findings.

The highest *rate* of duplication in BCL3 was in volume 3 (history) with 11.83% held by three or four libraries, followed by volume 2 (literature) with 9.22% held by three or four libraries; the least duplication was in volume 5 (science and technology) with only 5.88% held by three or four libraries. Volume 5 was also the volume with the greatest number of titles not held by any of the four libraries—29.41%. This probably reflects the tendency of materials in history and literature to remain undated and not to be superseded as compared to some items in technology and science, the greater reliance of the latter on serial literature than monographic, and the importance of many basic texts in history and literature to forming a core collection for the liberal arts curriculum. Clearly Rose-Hulman, being primarily a technical school, emphasizes different materials than do the other three libraries. Sanders *et al.* point out the need to build comparative data on disciplines; for example a library buying 30% of the publications in one discipline may be building a strong collection but in another discipline this might not be sufficient.¹⁴ For determining relative strengths within a collection as well as for comparing collections, this kind of data would be invaluable.

In the shelflist sample of LC classes USI had 31.4% as many titles as ISU and 10.4% of the combined total matched, whereas in the BCL3 sample the match (for ISU and USI only) is 25.8%. The BCL3

rate of overlap is predictably higher because it is an approved list and each school had equal opportunity to select those titles, based on budget and need; the LC classes are more varied due to the history of the collections, gifts, individual selectors, faculty requests, and other local factors. Chart 5 shows where among the LC classes the greatest overlap of the shelflist sample falls. For BC the overlap is 10.3%; for BH it is 18.4%; for HA, 14.5%; for LD, 12.1%; for CC, 9.4%; for PC, 5.1%; and for QM, 7.1%. For the first four LC classes the number that overlap is fairly high for USI, though less so for the remaining three classes.

A study of overlap in the eleven University of Wisconsin libraries found a lower rate of duplication than expected. "The actual rate of duplication falls between 18 and 32 percent for the entire UW System." The same source refers to several other studies showing overlap from 13% to as high as 76%, perhaps indicating variety in methods of study as well as variety in collections.¹⁵ Our findings ranged from 10.4% - 34.2%. Our numbers, though the range is slightly wider, are close. According to Potter, "the extent of collection overlap is dependent upon the age, size, and type of libraries involved in that the probability of overlap increases with size of library, but decreases when the libraries compared are of different age or type." The percentage of unique titles ranges from 50-86%, showing that "even the smallest library in a group has something to offer to a cooperative management." He notes another study that shows that it is "more productive to add libraries of different types or age to a union catalog project than to add large libraries of the same type."¹⁶

Conclusion

Given the sizes of our libraries, the numbers and percentages of overlap need not surprise. They may give us confidence to continue with current policies for the short term since relatively few volumes are duplicated. Since these findings are in line with past studies, we now need to move forward with cooperative arrangements. We are all aware of the increased ILL activity involved in resource sharing, but these changes will also affect reference and instruction, as students and faculty are taught to effectively use the greater resources available, as well as acquisitions and collection development.

An aspect of the study for the future is determining what kind of collection development and cooperation are needed. Those needs are viewed differently by state legislators, university presidents, library directors, collection development librarians, and patrons. As Stroyan points out, given the diversity found so far in overlap studies, "elaborate collection development programs to avoid duplication may not be necessary. Libraries are already doing this without formalizing it,"¹⁷ At the beginning of a cooperative venture, policy decisions are needed, but after the initial arrangements are made, perhaps meetings as needed and phone calls to maintain contact can suffice. Each library must be clear about its own mission and collecting strengths and communicate these clearly to others. For most of us that will continue to require much time and effort.

An interesting question to pursue at a later date is whether any of these overlapped items have been checked out and if so, how many times? Can those numbers/percentages be profitably compared

to the check out rate for other titles or LC sections? The shifting emphasis from collections to "access" brings into question many standard procedures, such as selecting from standard bibliographies, core lists, and indexes. To test whether that shift is working effectively would require some testing of the use of index/bibliography derived selections versus ILL supplied materials.

In some ways it is harder now with newly automated systems than it was before automation to obtain certain kinds of data on collections. The potentially greater amount of kinds of system-provided data are not yet readily available. The manual systems, despite their limitations, did give us some objective data; in many cases that record keeping was stopped when the library was automated. With the tremendous effort necessary to get accurate cataloging records into the database, there hasn't been time to replace those old manual records with improved online versions. One of the compelling needs now is for our automated systems to start supplying the kind of data upon which collection management decisions can be made—e.g., use data tied to LC classes, funding, and selector.

This initial study should enable us to ask appropriate questions and work toward answers. As we are able to extend this study, we hope to learn ways in which to improve our selection choices not only for our individual institutions but also for the network. The issues are important ones that will be with us for some time.

References

1. Medea Ionesco, "Regional Cooperation for Research Collections," *Collection Building* 9:2 (1989), 7, 9.
2. Thomas Nisonger, "Editing the RLG Conspectus to Analyze the OCLC Archival Tapes of Seventeen Texas Libraries," *Library Resources & Technical Services* 29 (Oct./Dec. 1985), 323.
3. David Farrell, "The North American Collections Inventory Project (NCIP): 1984 Phase II Results in Indiana," *Coordinating Cooperative Collection Development: A National Perspective*. ed. Wilson Luquire (New York: Haworth, 1986), 37-48.
4. Barbara B. Moran, *Academic Libraries: The Changing Knowledge Centers of Colleges and Universities*. ASHE-ERIC Higher Education Research Report No. 8. (Washington, D.C.: Association for the Study of Higher Education, 1984), 71.
5. Paul H. Mosher, "Cooperative Collection Development Equals Collaborative Interdependence," *Collection Building* 9:3-4 (1989), 30-31.
6. Nisonger, *op. cit.*, 309-327; Barbara Moore, Tamara J. Miller, and Don L. Tolliver, "Title Overlap: A Study of Duplication in the University of Wisconsin System Libraries," *College & Research Libraries* 43 (January 1982), 14-21; Nancy P. Sanders, Edward T. O'Neill, Stuart L. Weibel, "Automated Collection Analysis Using the OCLC and RLG Bibliographic Databases," *7 College & Research Libraries*, 49 (July 1988), 305-314.
7. Blaine H. Hall, *Collection Assessment Manual for College & University Libraries*, Phoenix: Oryx, 1985. "Appendix B. Statistical Aids," pp. 115-122 recommends 381 items for a population of 50,000, approximately the number of titles in BCL3; we chose to take a slightly larger sample.
8. Debora Shaw, "Overlap of Monographs in Public and Academic Libraries in Indiana," *Library &*

Information Science Research 7 (July 1985), 278.

9. This shelflist comparison was not extended to Rose-Hulman or St. Mary's, who use Dewey rather than LC, because it would have been necessary to check each ISU title against those holdings as well as checking USI's holdings against ISU's separately.

10. Michael K. Buckland, Anthony Hindle and Gregory P.M. Walker, "Methodological Problems in Assessing the Overlap Between Bibliographical Files and Library Holdings," *Information Processing & Management* 11 (1975), pp. 91-93.

11. Nisonger, *op. cit.*, 323,325.

12. Shaw, *op. cit.*, 283.

13. Shaw, *op. cit.*, 284.

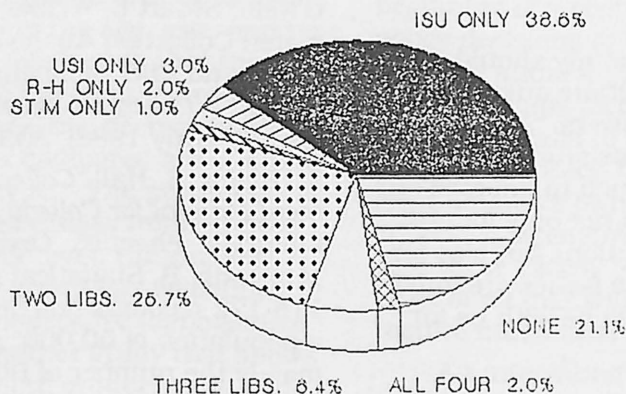
14. Sanders, *op. cit.*, 312.

15. Moore, Miller, and Tolliver, *op. cit.*, 20-21, 15

16. W. G. Potter, "Studies of Collection Overlap: A Literature Review," *Library Research* 4 (Spring 1982), 19 and 9, citing W.Y. Arms, "Duplication in Union Catalogs," *Journal of Documentation* 29 (1973), 373-379.

17. Sue Stroyan, "Collection Overlap in Hospital Health Sciences Libraries: A Case Study," *Bulletin of the Medical Library Association* 73 (October 1985), 358.

DISTRIBUTION OF BCL3 SAMPLE Unique & Overlap Holdings & Gaps

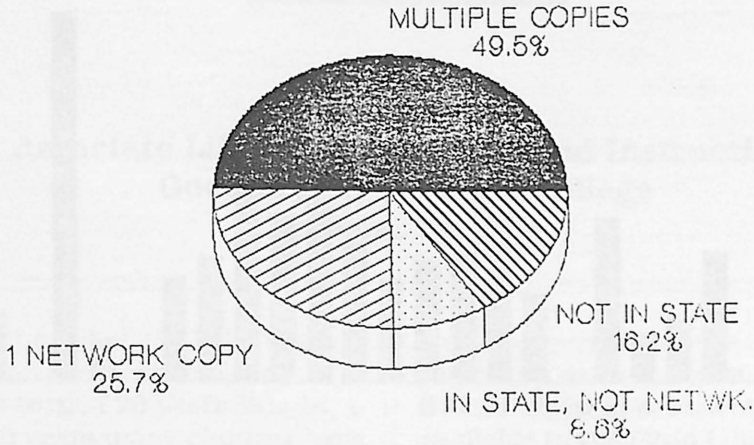


Based on 497 titles in sample

Chart 1

AVAILABILITY OF "GAP" TITLES

Percentages of Total "Gap" (105)

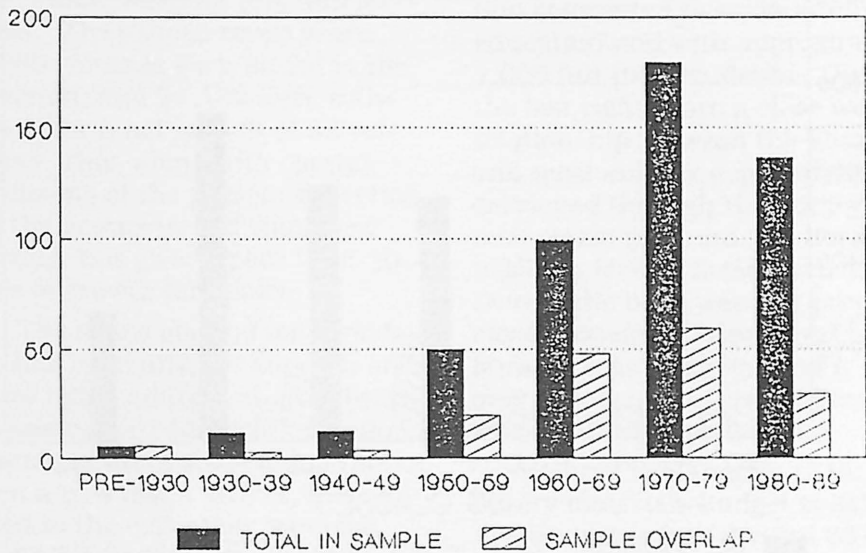


These represent 21.1% of total sample

Chart 2

BCL3 SAMPLE BY DECADES

Sample Totals and Overlap Holdings



Numbers Given Are Not Percentages

Chart 3

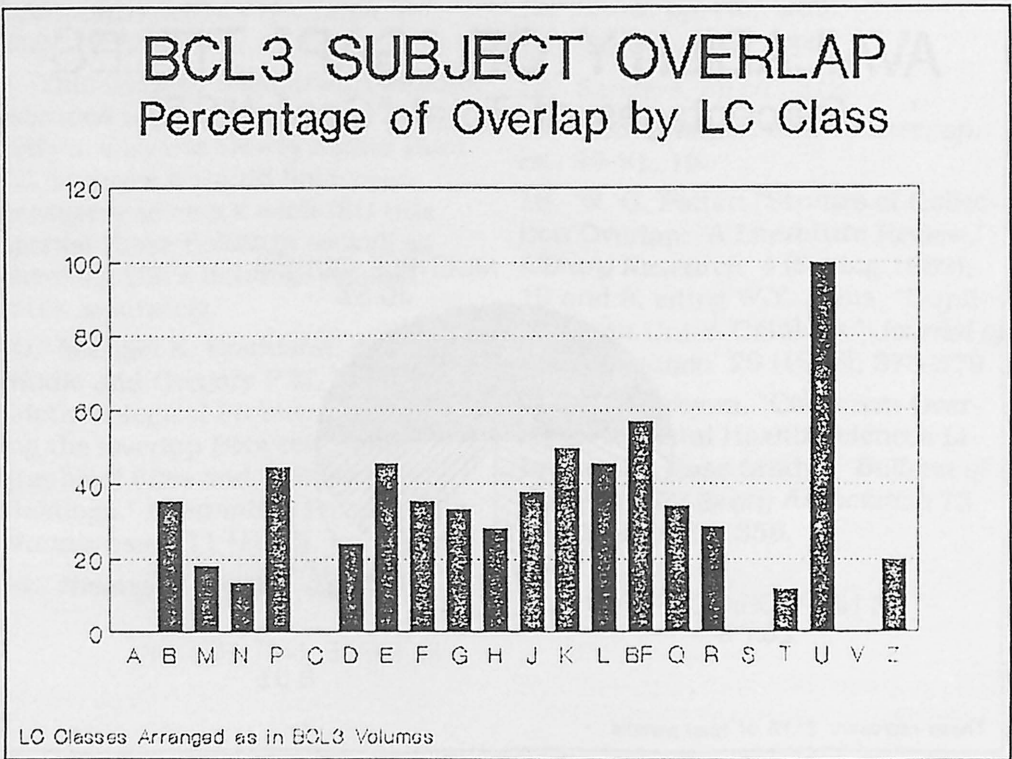


Chart 4

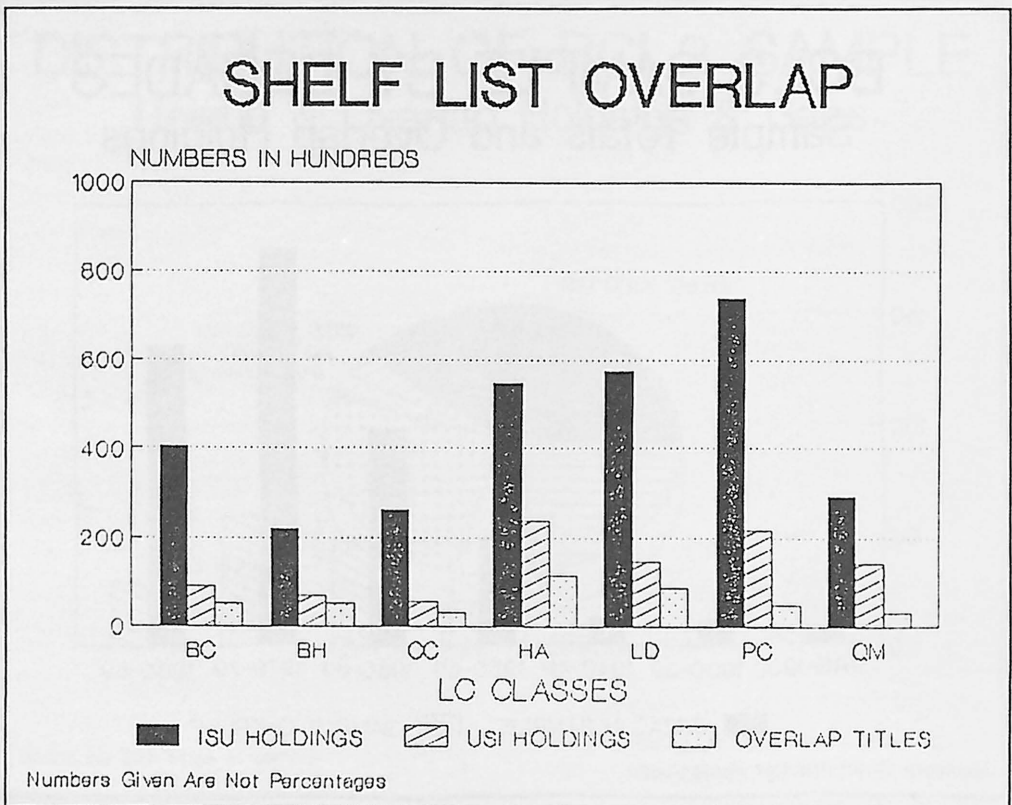


Chart 5