Collection mapping
- An Evolving Tool for Better Resources and Better Access

Mary C. Bushing

Collection evaluation techniques have evolved as a result of technological changes in libraries but their purpose and benefits have remained the same: the ability to understand the specific strengths and weaknesses of information resources with statistical data as well as impressionistic judgments based on experience and knowledge of the discipline area under consideration.

Librarians have adjusted the tools to evaluate the state of information resources in specific subject and format collections according to the technology available. Whatever the tool, the "picture" of resources that results enables libraries to adjust their holdings to better meet their missions, while also enabling us to streamline access to information without being tied to strictly bibliographic tags with subject and discipline gateways. Whatever the tool, however, its usefulness is dependent upon the thoughtful and appropriate application by the practitioner.

The concept of collection mapping or collection assessment is neither new nor is it particularly innovative as a professional library tool. The idea can be traced through the library literature and practice from the 1940's onward. It has been referred to by a variety of names but both "mapping" and "assessment" are the names most often used today.

The techniques and tools, like the terms themselves, have varied across time. As technology has advanced and its applications in libraries have become ever more complex. Libraries have been able to conduct assessments with increasing accuracy while at the same time reducing the amount of hours devoted to actually handling the collections themselves.

With almost forty years of broad experience by all types of libraries, the knowledge base for collection mapping or assessment now provides us with many options to select from as we plan assessment projects. This article will address the theoretical and practical history of this tool for informed library management. It reflects personal knowledge developed across thirty years of close involvement with many aspects of collection assessment and with projects in a number of countries.

Concept of collection mapping

It is important from the start to understand what collection mapping is and what it is not so as to put the entire process and its product into the correct context. Collection mapping is not:

- Exclusively a quantitative process like most title overlap studies that list titles and percentages of titles held in common with other libraries as well as creating title lists of unique holdings.
- Just data from bibliographic systems arranged in a special format.
- A simple linear process.
- A value judgment about the quality of a given collection.
- A rigidly defined set of actions, measurements, and techniques.

Rather collection mapping provides a broad range of operations and techniques to be selected to achieve the degree of informed collection understanding necessary in any given circumstances. It other words, collection mapping is:

- A logical outgrowth of thoughtful librarians attempting to see and understand their collections in subject segments not collection size.
- A tool made both efficient and more complex by technology.
- A flexible tool that requires continuous decision-making and clarity of definitions and purpose for any given project.
- A means of providing a snapshot of existing resources but which leaves to the library management the determination of the extent to which existing resources are appropriate for the given clientele.
- A tool that is never perfect-but is useful for comparison and planning purposes.
- A process that provides statistical information and defines the broad character of a collection of information and literary resources but requires many informed human judgments if it is to provide useful results.
- An evolving tool-not a solution or an end product but a means to an end.
- A tool that enables libraries to graph collection strengths across disciplines and/or subjects in order to justify expenditures in light of user needs and library missions.
- A tool to assist selectors in shaping collections to meet defined collection goals.
- A useful and functional tool for large and small libraries of all types.

In other words, it is a flexible tool that can be adjusted to fit circumstances within the parameters of a given project and its purposes. It can be used across languages and cultures to represent collection strengths and weaknesses in a meaningful way for funding agencies, librarians, researchers, and other library clients.

A collection mapping project can be defined within the constraints imposed by available resources-time, personnel, and technology. The extent to which detail is necessary to fulfill defined outcomes can also influence the resources necessary to complete a useful assessment project.

Collection assessment or mapping grew out of the need to be able to describe collections in ways other than title lists. The so-called information explosion helped to make this type of tool a necessity. As book collections and other information resources outgrew our ability to "know" them in any meaningful manner especially within large libraries, librarians began to consider how to better define the depth and breadth of their collections. Collection assessment was the result.

Learning the process

One of the things we have said in teaching others about the preparation and importance of writing meaningful collection development policies is that if done well, the process itself is as valuable as the resulting product. This is also the case with collection mapping or assessment.

While technology has enabled us to automate some of the baseline data gathering, assessment remains more than the data. If it is to be meaningful, the process must be constantly monitored and adjusted by informed staff members who understand the unique knowledge management issues within each discipline.

Staff members need to be able to interpret the data in a meaningful manner against the broad spectrum of human knowledge in all formats. Numbers are not the essence of the information. Statistics are only the first pixels in the map that will eventually provide a complete landscape of the information resources of a given library or group of libraries.

Prior to the evolution of this tool, libraries typically described their collections in quantifiable and statistical measures only. This type of information tended to imply that the more items in the library, the better the library. Working with libraries of all types, it became clear to me early in my career that this culture of "bigger is better" often led to poor collection management practices and collections that did not fit the needs of their clients.

This concept was further enforced by the accreditation process in academic institutions in the United States where secondary schools, colleges, and universities began to be required to have specific numbers of books per student. This discouraged librarians from actively managing their collections except in terms of acquiring more items.

The National Shelflist Count

The National Shelflist Count1 (later changed to the NATC or North American Title Count) or the North American Collections Inventory Project (NCIP) sponsored by RLG, were projects for large research institutions, and they further emphasized the importance of gross numbers of items.

With the National Shelflist Count however, the number of volumes were reported based upon classification ranges and thus basically subject or discipline specific items rather than total collection size began to be the standard. At this point no evaluation or professional judgments were used to provide any measure of quality for any given collection segment. Use statistics were used as the measure of the quality of the collection relative to the primary user group.

The National Shelflist Count was a major step in the development of what came to be called the "conspectus" method of collection mapping or assessment. A group of collection development librarians and members of the Research Library Group (RLG) began to be concerned about the failure of the statistics to indicate the character of collections, the academic level of the materials, or the emphasis of the collections reported.

Their concerns and discussions resulted in a series of committees within (RLG) devoted to the huge task of finding a means by which to organize or survey collections by broad academic subjects or disciplines using techniques that would better convey the character of collection segments with a structure that would reflect the typical university organization around modern disciplines rather than the Nineteenth Century world view that is reflected in both the Library of Congress (LC) and Dewey Decimal Classification (DDC) schemes.

The RLG Conspectus

In order to accomplish this, a model formerly used at the New York Public Library and Columbia University Library was expanded and used to develop what came to be known as the RLG Conspectus(c) with twenty-five broad divisions representing academic disciplines2. These initially identified divisions were subsequently expand-

ed to include additional divisions were areas such as Women's Studies.

Academic catalogers arranged LC classification numbers within these disciplines as an aid to librarians doing conspectus work. Thus worksheets were developed that provided the structure within which to gather statistical and other information with the goal of being able to define the character of the collection segment.

The RLG project worked towards a process that would encourage libraries to use a variety of methods to define the character of their collections and would take some account of the varied information cultures across discipline groups such as the journal driven literature of the sciences and the monograph nature of much of the humanities. In addition, they wanted the results to allow for meaningful comparisons across libraries. The name "conspectus" is actually a word for survey and aptly describes the concept of collection assessment by subject or discipline.

Co-operation between libraries

At about this same time, school librarians were developing a method to represent the character of their collections by gathering data and graphing that data within curriculum clusters to present a picture of their collections relative to the curriculum, funding levels, and educational goals of their library media centers.

The RLG group also wanted to be able to graph the character of the university collections along some type of well defined continuum. Public librarians in the United States were interested in finding similar means to define very popular areas of the collections such as adult fiction and children's materials.

Working independently but with their work gradually informed by the progress of the others, these various groups ultimately developed techniques and structures that were similar although targeted to the typical academic, public, or school library situation. The primary element that these various schemes had in common was that they focused on subject areas rather than whole collections.

They all required some basic statistical descriptors about each subject segment. These were: the number of items in each format within the subject, the average age of these items in terms of intellectual content, and the primary languages represented in the content of the subject segment. Additionally, most libraries involved in structuring this type of collection analysis saw the need for librarians to actually browse the materials in their subject area before being able to define the character of each subject segment.

The challenge of finding ways to easily convey the character of a subject collection was very difficult. Despite the expectation that individual libraries might have notes of local interest about each segment they examined, there was a need for some sort of shorthand that could quickly convey to others the level of scholarship in terms of likely users and/or audience level most likely suited for this subject segment. This presented special challenges since any type of code seemed to be fraught with unwanted evaluative connotations.

1-5 scale for collection definition

Ultimately, the RLG group settled upon a five-point scale simply using the numbers 1 to 5 and assigning each a definition that corresponded to levels of education and types of use understood by university librarians and university administrations.

For them a "1" collection was considered minimal while a "2" level collection could serve a general population of educated citizens and students through perhaps the first two years of college. They defined a level "3" collection as one that would support the work of college students in their major field of work and into their initial post graduate work for a Masters degree. A level "4" collection was one that could support doctoral study and other independent high level research and a level "5" was defined as comprehensive.

For the purposes of the large research libraries this scheme was quite workable and easily understood. Furthermore, it was possible to use this code system to relate discipline collections to the actual degree work offered at an institution and to identify both areas of strength congruent with the intellectual work being conducted in the institution.

Accordingly it was possible to identify those areas of information resources that needed to be strengthened in order to support the level of intellectual inquiry expected by the educational programs. The codes were used to define the level of existing collections, the rate of acquisitions within the same subject segment, and the goal level for that segment.

Other techniques

Those working in public libraries initially wanted a coding system that would be based upon an alphabetical code such as A, B, C, etc. This system was rife with value problems also because these same letters are those used as grades in the educational system and imply a judgment that was not quite appropriate for what the librarians wished to convey. For all of those working on the collection mapping idea, they wanted their work to reflect a description of the collection segments surveyed but not a value judgment.

Thus an "A" or an "F" collection of popular fiction (depending upon which way they determined to use their alphabetical code) did not mean that this was the best possible collection but a collection that contained certain elements and was likely to be quite large. It was not meant as the ideal towards which other libraries should strive but it was hard not to think of an "A" collection as the ideal.

The school librarians avoided the problems associated with coding the collection by devising various schemes by which to use bar graphs to convey the extent of curricular area collections with additional statistics reported either as contrasting graphs or in textual explanations. This approach worked quite well for their purposes within, rather than across, school buildings and districts.

Gradually, other techniques were recognized as helpful in the process of mapping collections. The most widespread techniques were comparisons with other similar or ideal library collections, comparisons with standard lists and bibliographies as well as journal titles indexed in the standard periodical indexes within a discipline, and another type of list checking, citation analysis.

These techniques along with shelf scanning and the statistical elements already noted and shelf scanning came to be the most often used techniques. Additionally, the use of circulation or use statistics were used to help inform the collection picture as librarians wished to move towards more judgmental positions regarding the collection pictures that were forming from the assessment work.

A snapshot of the mathematics collection was useful but adding to that snapshot the statistics regarding the actual use of the mathematics collection, the amount of money being expended on that collection annually, and ultimately, the purpose or goal for the mathematics collection in light of the mission and potential users of the library provided a more complete picture of the mathematic resources and helped to define appropriate goals.

Some libraries also sought statistical information about the amount of materials that they needed to borrow from other libraries, if any, in support of their users' need for mathematical information. The tool of collection assessment or mapping continued to morph as librarians used it and sought ways of defining their resources in relationship to their clients and their missions.

Linkage to LC and Dewey Decimal Classification

During this period of the 1970's, the United States' academic community continued to develop their schema for conspectus based upon

broad discipline areas and using LC classification as the keys to actually identifying those collection segments in support of a particular discipline area.

Their schema admittedly was very U.S. centric due to its basis in the Library of Congress Classification outline and the fact that those involved were all from academic institutions in the United States. There was nothing similar for those libraries using Dewey Decimal Classification, however.

In the early 1980, some uniformity was established due to the work of a group of outstanding visionary librarians in the state of Alaska who understood the applications of assessment for cooperative collection development even within library communities as far flung as those in the wilds of Alaska.3

With their leadership, grant funding from what was then the Fred Meyer Charitable Trust,4 and the participation of librarians from across the Pacific Northwest region, the RLG conspectus was adapted and simplified with a DDC focus and with new collection level definitions that would provide more precise definitions of collection characteristics for public libraries, smaller academic libraries, and even school libraries.

After lengthy discussions and some hand wringing, the resulting codes for the new conspectus approach including not a five-point scale but a ten-point scale that incorporated the five-point RLG scale so that comparisons were still possible. The new scale (0 - 1a - 1b - 2a - 2b - 3a - 3b - 3c - 4 - 5) became the standard and was adopted by those previously attempting to use an alphabetical scale.

In addition, codes were developed and revised over time to include language codes that easily accommodated situations were countries have dual languages, the inclusion of electronic resources in the expanded definition for the codes, and even a standard scale that could be used when appropriate to indicate preservation issues. A summary of the basic coding system appears in Table 1.5

Impacts of library automation

While the practice of collection assessment or mapping was growing during the 1980's, the labor intensive aspects of gathering data were daunting for many. Despite the work of collection librarians working with library automation vendors, we were still unable to get title counts by specific classification ranges, average publication dates within those same ranges, and other pertinent data.

Additionally, although software in support of the conspectus (both LC and Dewey) had been developed, the software was initially focused on two aspects of conspectus: printing standardized worksheets and printing reports although the ability to present those reports in graphical form was a huge plus that made communicating the assessment information to those outside of the library very easy and effective.

As the 1990's unrolled, however, library automation had advanced to a point where it was easier and easier for technology to come to the rescue of those wishing to map their collections and the specific characteristics of each subject segment. It became possible to get more accurate title counts than those achieved through estimation from the stacks or shelflist measurements.

Additionally, it became possible to get statistical reports giving the mean age of collection segments based upon copyright or publication dates. Circulation reports could also be narrowed based on classification ranges.

OCLC and other bibliographic vendors began to offer comparison reports against other collections defined as peers or ideal collections or against standard lists such as Books for College Libraries.6 The software in support of the Pacific Northwest project was moved to WLN (at one time the Western Library Network and later incorporated into OCLC) and the software was further developed to incorporate unlimited notes fields that were keyword searchable.

WLN also developed, although in FoxPro, a software platform no longer supported, the ability to provide year-to-year comparisons of statistics, progress towards goals, expenditures, numbers of items, mean age, etc., thus enabling a single library to easily track their progress and the changes they were attempting to make with their collections.7

Benefits of conspectus

Thus over the course of forty years, the library profession had moved from a general concept of trying to describe collections in more meaningful terms to express the character of the subject segments to a rather sophisticated level of descriptive reporting.

Assessment or collection mapping was useful in many political and financial arenas as well as within the library community itself. It enabled libraries to be responsive to demands by funding sources to justify increased budgets with data, graphs, and rationale that directly related missions to the existing resources.

The conspectus information enabled consortia and groups of libraries to compare resources and to develop clearly defined means of sharing resources of strength. It enabled individual libraries to identify "holes" or problem areas of their collections and to take action to remedy those situations. And lastly but not least, the use of collection mapping provided the essence of information to inform collection policies.

One of the most valuable outcomes from the conspectus development was the realization, that like developing collection policies, the process of doing collection assessment was perhaps as valuable as the results themselves. The librarians involved learned that doing the collection mapping resulted in greatly expanded knowledge of the collections both in terms of general character-

istics but also in terms of very specific strengths and weaknesses.

In addition, the process provided a forum in which it was possible to share a wide-range of knowledge about the world of books, authors, publishers, discipline specific communication cultures and history, and the nature of differing types of library collections.

During the stages of collection mapping these benefits were evident in the following ways:

- Discussing the appropriate data to gather and then gathering it;
- Defining and understanding the differences in information cultures from discipline to discipline, from subject to subject;
- Deciding upon appropriately useful levels of definition for segments or subjects;
- Identifying consistent vocabulary, codes, and notes; and
- Arguing issues such as goals and descriptive codes with subjective although not uninformed opinion as part of the process.

Perhaps one of the most valuable aspects of the experience from the late 1980's and early 1990's was the realization of just how very adaptable for every size and type of collection the methodologies can be. Projects in national libraries in Australia, New Zealand, Latvia, and the Czech Republic among others testify to the ability to adjust the process and the product to fit very individual circumstances.

The recommendation by both ALECTS (an American Library Association division, the Association for Library Collections and Technical Services) and the International Federation of Library Associations (IFLA) for conspectus methodology to be used in describing collection strengths and their characters have continued to encourage libraries to adapt these assessment methodologies and structures for their collections. In many areas of the U.S. collection mapping projects within individual libraries are based upon the extensive experienced gained during the 1980's and 1990's.

The end results and primary benefits from using collection mapping are the extensive knowledge of collection details and the nature or character of the existing collections. This allows libraries to develop a clarity of purpose that has often escaped them in the past.

It gives staff and administrators an opportunity to define real problems to be resolved with regards to information resources rather than merely spending more money. It enables librarians to document policy decisions and collection goals. It gives staff an in-depth knowledge of discipline specific information cultures and gives them confidence in making day-to-day decisions regarding information resources with more focused acquisition goals.

The global scale

The situation now is that World Cat and the OCLC database have not only grown to over 9,000 members and over 65 million records with over a billion items,8 but the focus is increasingly on end-user searching, sharing & comparisons on a global scale. Metadata, new protocols, and standards are now enabling collections themselves to be described and accessed through records that reflect the collection not the items within it.

Subject gateways that are easily navigated and identified are being made possible through a wide range of software. These developments are enabling the integration of collection mapping results in existing electronic bibliographic databases and search tools.

It is an exciting time in the development of the tool of collection assessment. For someone who has been closely involved in the evolution of this tool over the past thirty-five years, it seems to be the realization of the goal of assessment: the ability to inform users seamlessly about the character of subject or discipline collections throughout the world while enabling libraries to better collaborate in building upon strengths and strengthening weaknesses where needed.

To identify collections of note, it still requires

the work of collection mapping to identify the strong subject collections for they are not always the largest! It means that data gathering, analysis, and informed decision-making remain key to using subject gateways and metadata in meaningful ways.

Lessons learned

Before concluding this survey of the history and lessons we have learned with the evolution of the assessment process, it is essential that potential users of this tool, in whatever form, recognize essential characteristics for the project leaders and other individuals involved in a collection mapping process.

From experiences we have learned that those involved must:

- Be comfortable with a reasonable degree of ambiguity but have analytical skills;
- Be able to be dispassionate regarding their own collections;
- Have facilitation skills to bring others to consensus during all phases of the process;
- Have project management and cheerleading skills;
- Have the ability to make decisions and not dither;
- Have an understanding of the differing information cultures discipline to discipline;
- Have clarity of purpose with an understanding that the process and the results are not the ultimate goal; but,
- Understand that the ultimate goal is to improve appropriate resources and to provide better access to those resources.

These lessons regarding the type of personalities and skills essential for collection mapping were learned the hard way-through experience. They were learned by being involved in all types of libraries, of differing sizes, in many different countries. Collection assessment often was centered in technical service departments initially but projects became bogged down in the data gathering possibilities. Projects centered in public service and reference departments often had a hard time actually generating the necessary data and using it.

The choice of a project leader in all cases was a key element for success. Leadership, credibility, knowledge, and an appreciation of both the necessary quantitative elements as well as the qualitative, judgmental elements that are needed to appropriately describe the character of a collection are essential personal qualities for a collection mapping project leader within a library.

In 2006 we find ourselves in an information environment where there are a variety of approaches to collection mapping. While it is possible to merely gather statistical data or even just do title over-lap studies, those do not result in increased competence on the part of the library staff nor do they, in my experience, do anything other than result in long reports that are not much more helpful than the shelflist counts of the past.

These approaches return us to the quantity not quality or character descriptions of collections. Consider the difference between having 400 books about dinosaurs written for children and the 400 volumes held a few hundred feet from my study window by the Museum of the Rockies where internationally renowned paleontologists like Jack Horner do their work.

Much of the statistics about these collections could be the same, number, age, format, even titles, and use statistics. What is glaringly different is the character of these collections. The one appropriate for individuals aged perhaps 4 to 12 and the other appropriate for informed, educated, specialists doing research, digging for, and reconstructing an extinct group of animals.

The bottom line is: with or without the aid of technology, librarians find it helpful to map their collections to better understand their existing information resources and to define appropriate goals to improve both the quality and access to the most appropriate resources given a library's mission and its current and future clients. The process and the product are valuable on the

local level as well as valuable on a national and/or consortia level for collaborative information resource development and access.

The five or ten-point codes developed for describing collections still serve well and can enable a library or group of libraries to succinctly convey the nature or character of a collection segment. They are not perfect, but then what is!

Doing something, however imperfect, is still better than allowing the perfect to become the enemy of the good. The evolving tool of collection assessment or mapping is not perfect but it is what we have and it is a great deal better than it once was and better than any of the alternatives as of yet suggested.

TABLE 1

0		Does not collect intentionally
1	a	Minimal, uneven, unsystematic
	b	Minimal, focused coverage, consistently maintained
2	a	Basic information level, introduce & define a subject, basic general monographs, some subject periodicals, basic reference tools
	b	Broader and more in-depth array that include history of the discipline/subject, important personages, broader array of reference sources, indexes & electronic resources
3	a	Basic study/instructional support, high percentage of most important sources, core works, extensive collection of periodicals in the subject, access to appropriate electronic sources, works in primary language of the clients, undergraduate materials
	b	Intermediate study/instructional support, more specialized subject areas, more comprehensive coverage, high percentage of core works, well-known authors in their original languages, specialized resources in all formats including electronic, larger, more in-depth collection across most aspects of the discipline, supports upper level undergraduate study and initial post-graduate study
	c	Advanced study/instructional support, resources for imparting and maintaining knowledge about all aspects of the topic, a large collection, many works in other languages, primary material & extensive secondary material, lesser known as well as core authors, supports master's level & doctoral course work
4		Research level, older material is retained & systematically preserved, access to extensive runs of all key journals, reference sources, & monographs, supports doctoral study & independent scholarly research.
5		Comprehensive/exhaustive level, strive to be exhaustive as far is possible (i.e., special collections), extensive manuscript collections, extensive collections in all pertinent formats, exhaustive published materials in many languages, for historical research.

1 The first National Shelflist Count was held in 1973 with 17 institutions participating. In 1993, as the North American Title Count, data was collected from 56 libraries. Yelverton, John. (1997, April 21). The 1997 North American Title Count: Information and a Call for Participants. Message posted to COLLDV-L electronic mailing list, achieved at http://www.usc.edu/COLLDV-L

2 Agriculture, Anthropology, Art & Architecture, Asian Studies, Biological Sciences, Business & Economics, Chemistry, Computer Science, Education, Engineering & Technology, Geography & Earth Science, History & Auxiliary Sciences, Language, Linguistics, & Literature, Law, Library Science, Mathematics, Medicine, Music, Performing Arts, Philosophy & Religion, Physical Education & Recreation, Physical Sciences, Political Science, Psychology, Sociology.

3 Stephens, Dennis. "A Stitch in Time: The Alaska Cooperative Collection Development Project." In Coordinating Cooperative Collection Development: A National Perspective, edited by Wilson Luquire, pp. 173-84. New York: Haworth press, 1986. Also published in Resource Sharing and Information Networks 2 (1985).

4 The Fred Meyer Charitable Trust provided funding for the first few years of the Library and Information Resources for the Northwest (LIRN) Project. It published the first Pacific Northwest Collection Assessment Manual in 1985 and later the project was funded by both the Fred Meyer Trust and the state libraries of Alaska, Idaho, Montana, Oregon, and Washington. The LIRN Project obtained the rights to the RLG Conspectus and was able to use it as a model in their work. See: Forcier, Peggy, "Building Collections Together: The Pacific Northwest Conspectus." Library Journal 113 (April 15, 1988): 43-45.

5 Bushing, M., B. Davis, and N. Powell. (1997). Using the Conspectus Method: A Collection Assessment Handbook. Lacey, WA: WLN, pp. 27-31..

6 Books for College Libraries: A Core Collection of 50,000 Titles. (3rd ed.). (1988). Chicago: American Library Association.

7 Pinnell-Stephens, J. (1992, November-December). A management information system using the WLN Conspectus software. WLN Participant, 11(6), 12-14.

8 OCLC webpage at http://www.oclc.org/worldcat/default.htm retrieved April 13, 2006.

Mary Bushing, Ed.D. Professor Emerita, Montana State University-Bozeman, Bozeman, Montana U.S.A.
Currently, Library Consultant & Educator
email. marrying@msn.com