What Is a Collection?

Hur-Li Lee
School of Library and Information Science, University of Wisconsin–Milwaukee, P.O. Box 413, Milwaukee, WI 53201-0413. E-mail: hurli@uwm.edu

Advances in information technology have dramatically changed information seeking, and necessitate an examination of traditional conceptions of library collection. This article addresses the task and reveals four major presumptions associated with collections: tangibility, ownership, a user community, and an integrated retrieval mechanism. Some of these presumptions have served only to perpetuate misconceptions of collection. Others seem to have become more relevant in the current information environment. The emergence of nontraditional media, such as the World Wide Web (WWW), poses two specific challenges: to question the necessity of finite collections, and contest the boundaries of a collection. A critical analysis of these issues results in a proposal for an expanded concept of collection that considers the perspectives of both the user and the collection developer, invites rigorous user-centered research, and looks at the collection as an information-seeking context.

The Internet has revolutionized the way people access information. In working with growing numbers of electronic information resources in addition to traditional materials, information professionals are in the process of reexamining their roles and work practices. One concept that has become increasingly problematic is "collection."

There are many intriguing issues related to the conceptualization, or reconceptualization, of the library collection. For instance, can a collection exist in a virtual form? If the idea of a virtual collection is accepted (e.g., a digital library collection), does it matter where the digital files are located? If it is a collection, where are its boundaries? When a World Wide Web (WWW) page has hyperlinks to other WWW pages, is it a collection or a directory? Do other sites linked to the selected sites on this page also belong to the collection? If they do, does this literally mean that all digital information resources that are linked to one another, in one way or another, form one gigantic collection? Does this, then, mean that the idea of "collection" is meaningless in the digital age, especially to users? Finally, can a collection be developed automatically by a computer, and if yes, are there any differences between computer- and human-developed collections?

These and other related questions have a significant bearing on information seeking as well as library collection development. A major function of a traditional library collection is to facilitate information seeking by providing its users with convenient access to relevant information resources (Buckland, 1992). As more information becomes accessible electronically, this function of collection needs a fresh examination. We now should ask what access means in this context, especially from the user's point of view. Now more than ever, it is crucial for us to better understand collections and how collections facilitate information seeking.

Undoubtedly, the collector's understanding of a collection is what will determine the nature and scope of a collection and the way in which the collection is developed, maintained, and evaluated. As information services diversify, many more information professionals than librarians are getting involved in developing collections of information resources. A useful concept of collection will not only help librarians refocus their collection efforts but also provide others with valuable guidelines for designing new information services.

This article aspires to reconceptualize collection by concentrating on functional aspects of this concept in light of recent technological developments. To that end, the next section critically reviews some traditional characteristics that have been associated with a collection. In the third section, problems that render traditional concepts of collection unsuitable will be examined. The final section proposes an expanded concept of collection and explores its implications.

The term "collection" has many meanings, depending on the context of its use. For example, an anthology is frequently referred to as "a collection of works." In the archival world, a collection consists of a group of documents originating from the same source and acquired as a whole. The following discussion concerns neither of these two situations. Instead, it centers on the issues germane to the function of collection development in information services. In other words, "collection" should be thought of as in the term "a library collection"—an accumulation of information resources developed by information professionals intended for a user community or a set of communities.

Traditional Conceptions of "Collection"

Most popular textbooks on collection development offer no formal definition of collection (Curley & Broderick, 1985; Evans, 1995; Gardner, 1981; Spiller, 1986). Nor is a search in the library and information science literature particularly fruitful. The growing literature on digital libraries (e.g., Association of Research Libraries, 1995; Kuny & Cleveland, 1998; Saffady, 1995) has also left the concept of collection undefined. On the surface, this may imply that "collection" needs no definition, or that everyone knows exactly what a collection is. Only when one looks more closely do discrepancies become discernable. Some people perceive library collections as mere aggregates of physical packages of information. A narrower view limits these physical packages to print and text-based sources such as traditional books and periodicals. On the other hand, most librarians, as well as skillful users, have a more inclusive concept of collection.

Let us first examine two of the definitions of collection available in the literature. The first is a formal definition given in the *Encyclopedia of Library and Information Science* (Kent & Lancour, 1971),

> A library collection is the sum total of library materials—books, manuscripts, serials, government publications, pamphlets, catalogs, reports, recordings, microfilm reels, micro cards and microfiche, punched cards, computer tapes, etc.—that make up the holdings of a particular library. (v. 5, p. 260)

Two major presumptions in this definition have significant implications: tangibility[1] and ownership.[2] In listing types of materials in a collection, this definition seems to include only those that are tangible. This is hardly surprising, given the work was published in 1971, before the proliferation of virtual information resources. Second, the word "holdings" explicitly connotes the idea of ownership by the library. These two presumptions have greatly shaped the management of traditional library collections. In the library operations research literature, for example, authors often are concerned about the size and maintenance of holdings (Lee, 1993; Trueswell, 1969). Their operational definition of a collection relies heavily on the library statistics collected by following standard manuals (e.g., American Library Association, Statistics Coordinating Project, 1966) that mostly center on counting physical volumes owned by the library.

The second definition is more recent (Peek, 1998) and states,

> Libraries [in the past] were a collection of information, usually databases called books, held in a specific location. The presumption is that this information was intended to be shared—perhaps not shared with the entire world, but available to a specific community . . .Walls were a practical concept . . . and libraries owned the information contained within the walls. (p. 36)

Peek apparently interprets a library as a collection. Nevertheless, the above quote may be regarded as a definition of a traditional collection. In addition to the two presumptions mentioned in the first definition, this second definition offers a third one: a user community.

These three essential presumptions often dominate traditional concepts of collection and warrant further discussion. Further, a less documented concept of viewing a collection through a retrieval system, such as the online public access catalog (OPAC), also needs to be considered and, therefore, is examined in this section as well.

Tangibility

Some people speculate that libraries as we know them will soon become dinosaurs because the need for local (physical) libraries will evaporate when all information is transmitted through electronic networks (Kurzweil, 1992). The assumption here is that the library is a physical warehouse storing only tangible documents, and thus, all concepts and entities associated with the traditional library, including a collection, connote tangibility (Lagoze & Fielding, 1998). This view is somewhat common among technology enthusiasts, and even some library users who have used virtual information resources in the library. The long history of the library being associated with a physical building may have resulted in this fixed impression and made imagining virtual collections difficult.

A review of the collection development literature, however, clearly shows that the library collection gradually has expanded its scope. The resources collected have gone from print materials (Haines, 1950; McColvin, 1925) to a wide variety of nonprint and electronic resources (Evans, 1995; Scholtz, 1989; White & Crawford, 1997). Though some librarians initially may have resisted new formats, many have now adopted them. At present, it is common for libraries of all types to collect more than just the traditional formats.[3] To librarians, the reasoning is simple: the library

[1] People, in general, consider electronic resources intangible even though these resources are stored in a tangible devise, for example, a remote computer server. Thus, a tangible item, in this article, means an item that has a local physical presence.

[2] Librarians generally contrast ownership with access, and consider that the library *owns* the physical information packages, for example, books that are purchased by or donated to the library. This article adopts this general usage. For the debate on ownership/access, see Gorman (1997) and Lee (1994). The legal definitions and related issues of ownership/access are intentionally omitted from this article.

[3] For example, the new guidelines for American school library media programs developed jointly by the American Association of School Librarians and the Association for Educational Communications and Technology (1998) states, "the school library media specialist provides access to a wide range of electronic and other nonprint resources as well as to their traditional counterparts" (p. 84).

should acquire any useful, appropriate, and affordable information resources, regardless of format.

A tricky question is whether remote information resources are part of the library collection, for they are not physically collocated there—they may be distributed on separate servers and accessible only through an electronic network. In librarianship, there have been a few examples that provide precedents of a collection without being physically collocated. One of them is the concept of a national collection.[4] Since 1988, Australian librarians have vigorously promoted the formation of a *Distributed National Collection* (Waters, 1992).

It seems unhelpful to debate whether or not a group of objects, tangible and/or virtual, physically collocated or distributed, qualify to be a collection on a purely ideological level. Collections are developed for the purpose of serving users' information needs. The conceptual understanding of a collection must fulfill this practical purpose. Thus, how users perceive a collection during information seeking, how developers do so during collection development, and how a concept of collection can facilitate information seeking are more pertinent considerations.

As more information becomes available in digitized formats, information services have increasingly collected intangible documents, in addition to the tangible ones. The adequacy of applying tangibility and physical collocation in defining collections is questionable. Research on both the user's and the collection developer's perspectives is needed to shed light on this issue.

Ownership

Some say that traditionally a collection implies ownership (Hill, Janee, Dolin, Frew, & Larsgaard, 1999). In this conception, remote resources are not owned by the library and are, therefore, not part of the collection. For example, through interlibrary loan (ILL) agreements, a library patron in Chicago can borrow a book from another library in New York. This book is not part of the Chicago library collection. The patron may access it, but the Chicago library does not own it. A virtual document in a remote database seems similar in that the remote document can be accessed by the library's patrons, but is not owned by the library.

The requirement of ownership is highly limiting for three reasons. First, hundreds of American public libraries have long-term agreements with vendors to lease paperback books that have passing high demand. These leased books are not owned by the library but are always considered by librarians and users to be part of the library collection (Lynch, 1981). Second, an ILL book and a remote document are handled differently by the librarian. A book borrowed through ILL is only intended for one use and it does not go through the collecting process. In contrast, a remote database, such as the *Library and Information Science Abstracts*, is a conscious selection by a collector. The process of including a database, and thus all documents in the database, in a collection involves many considerations, and once made available, the database and its documents can be accessed by users of the community for repeat uses. It is likely for the user that those remote documents accessible immediately are no different than the books available in the local collection. Third, many bibliographic databases are on CD-ROM. It does not seem useful to categorize a database as part of a collection when it is on CD-ROM but as external to the collection when it is accessed remotely. To users, the difference between the two is hardly noticeable.

Ownership versus access has been a recurrent debate in the library literature. As librarians increasingly advocate access over ownership, a new concept, "shared collections," has appeared (American Library Association, Reference and Adult Services Division, Collection Development Policies Committee, 1993). Shared collections belong to different libraries, and may be accessed by people served by these libraries through interlibrary cooperation. To some librarians, all tangible items that their users may borrow through ILL practically are part of their collections, even though these items are the property of other libraries (Gorman, 1997). This idea expands the boundaries of a local collection, with a few catches. ILL borrowers are at best second-class citizens. For example, they have no right to recall an item checked out to a primary user of the library that owns the item. Unequal access, a waiting period, and sometimes, the fees incurred often make ILL unattractive (Truesdell, 1994). Sharing documents may lower the boundaries of the collections being shared, but does not eliminate the boundaries completely. Ownership is usually what establishes the boundaries in the first place.

The ownership issue needs special scrutiny in increasingly complex environments. We no longer see a singular dichotomy of owned or not owned. Publishers are making information packages available in a number of ways: for sale, lease, on-demand material delivery, and remote access only. Information services have additional options to provide information through interinstitutional cooperation, such as ILL. It is possible that users and information professionals have somewhat different ideas in terms of ownership. Information professionals are mostly concerned with how to secure and control information resources as well as the legal ramifications of doing so. Users, on the other hand, care about access and convenience. Any new understanding of a collection and collection development must deliberately examine this difference.

A User Community

Contemporary libraries acknowledge the fact that they cannot afford to collect all information resources, and emphasize instead that collections must be developed for *use* by their current user communities, be they residential, aca-

[4] The definition of a national collection according to the International Federation of Library Associations & UNESCO (1977) reads: "the collection of library materials held in a country . . ." (p. 9).

demic, or corporate. The quality of information resources should not be the sole consideration in collection development. Librarians have learned valuable lessons over the years, and are convinced that an effective collection must be developed with a solid understanding of its community information needs (Curley & Broderick, 1985). A group of high-quality information resources that do not meet users' needs is not a useful collection.

Consider Yahoo! as an example. It was developed for all people, rather than for a particular group. It certainly fits the dictionary definition of collection—a group of objects. But, does it satisfy our professional definition of a well-developed, useful collection? When 10 year olds navigate in Yahoo! to find information on the American Civil War, they will retrieve hundreds of sites, not only those created for children but also those intended for American historians. Information seeking in Yahoo! is unnecessarily time-consuming and frustrating precisely because the collection lacks the value of selectivity and customization.[5]

Thus, to make a collection useful, the information professional must select individual items carefully, basing the decisions on the community's needs. There is little question that the intended user community should be a critical element of a good collection.

A Unified Retrieval Mechanism

In some cases, even in the traditional view, a collection is not limited to resources in one building. For example, a collection developed for a university community often is separated into a number of physically dispersed subcollections, such as for the social sciences, medicine, and engineering. Large metropolitan library systems often have numerous branch libraries. One may think that these subcollections actually are themselves collections. However, one may also argue that membership in a library system extends to the entire university or city community: there is one cohesive policy for collecting resources that are intended to be one collection, and thus, branch collections are often referred to as subcollections.

In such large library systems, users frequently perceive the collection through one unified retrieval mechanism: the union catalog. Hill et al. (1999), in describing traditional collections, state, "A library's catalog is the index to the library's collections and contains the metadata for the items in those collections" (p. 1169). Modern technology has reinforced this conception. Users now can use the OPAC (an on-line version of the union catalog) at any location, even home or office, to access physically separate subcollections. The kind of library arrangement that encompasses physically distributed subcollections reduces the browsability of the whole collection. Nevertheless, the collection has become more integrated, from both the use and administrative points of view.

[5] Yahoo! is a directory service on the World Wide Web at http://www.yahoo.com.

Using a retrieval mechanism such as the OPAC to view a collection creates another problem, because many information resources in a library collection are not accessible through the OPAC. Some resources were cataloged before the library installed the OPAC, but the library has yet to convert those records into the machine-readable format. Others have not yet been cataloged but may be retrieved for use upon request, for instance, items in the cataloging backlog. A few subcollections in the library such as vertical files have unique features, and librarians sometimes create other *ad hoc* retrieval mechanisms to manage them; such items still are part of the collection, though not represented in the OPAC. Recently, many libraries have begun to collect documents on the World Wide Web (WWW) in a fashion similar to that of vertical files. Often, the selected WWW resources are uncataloged. The library offers a resource page on the library's Web site to serve as an *ad hoc* retrieval system that contains hyperlinks pointing to the selected WWW resources. Thus, these WWW resources are included in the library's collection by virtue of their being linked on the library's Web site.

Unfortunately, the gaps in a library's OPAC cause a major hurdle for users. They frustrate users by making part of the collection inaccessible from the main entry point into the collection: the OPAC. It also burdens users by forcing them to switch among a number of different information retrieval systems (IRSs) to find all materials in a library's collection. Although it is desirable from the user's perspective to access all information items through an integrated IRS, this is not the case at present in many American libraries. In system design, information professionals—librarians in particular—need to take this consideration seriously. In other words, an integrated retrieval system should be an indispensable element of a well-developed collection.

Challenges Posed by Media

Before the 1950s, many librarians gave little consideration to audio–visual materials for library collections (Quinly, 1956), but today, it is no longer a question and most libraries regularly collect AV materials. The Internet and hypermedia are generating more concerns for collection developers than those in the past. Two of the most complex issues are the controversy of disintermediation and the difficulty in ascertaining the boundaries of a collection.

Disintermediation

As information technology advances, questions regarding the necessity of forming finite collections arise. Some propose that technology will soon be fast and sophisticated enough to make all information resources stored, organized, and accessed on-line, forming one undivided information universe, and users may choose any resources by themselves in this universe without the interference of human intermediaries—this phenomenon is called "disintermediation." To them, a finite collection of resources selected by

an intermediary, a collection developer, only imposes unnecessary restrictions that inconvenience information seeking. Or, does it?

Research indicates that users tend to base decisions in information seeking on their own benefit, cost, and/or effort (Hardy, 1982). It is generally agreed that saving money, time, and effort is important, and the question is: are users willing to sort through the information universe by themselves? In reality, this universe has become so enormous and the volume and variety of information so overwhelming that seeking information in this undivided universe is neither feasible nor economical. The problem of information overload is a major theme that has been tackled by many authors in many fields (Alesandrini, 1992; Lively, 1996; Wilson, 1996). To minimize cost and effort, people understandably appreciate a filter that can exclude irrelevant and substandard information resources but make available high quality and useful items.

Information service professionals have refuted the prophecy of disintermediation in the essential functions they perform (Atkinson, 1996). To them, intermediation is a process of adding values to information resources. Robert S. Taylor's research (1986) identifies 23 types of added values. Of those, a well-developed collection offers many, such as selectivity, comprehensiveness, currency, reliability, and response speed. Selectivity in Taylor's words is "the value added when choices are made at the input point of the system, choices based on the assumption of the appropriateness and merit of certain information chunks or data to the client population served" (p. 61). In other words, a collection may work as an effective filtering system that helps reduce information overload. No doubt, selectivity continues to be a desirable added value in an increasingly virtual information world, as testified in the working definition of digital library proposed by the members of the Digital Library Federation (U.S.), which states, "Digital libraries are organizations that provide the resources, including the specialized staff, to *select* [italics added] . . . collections of digital works so that they are readily and economically available for use by a defined community or set of communities" (Waters, 1998, p. 1).

A related issue is how intermediation is done. Lagoze and Fielding (1998) propose that a digital collection is "a set of criteria for selecting resources from the broader information space" and "provides tools for resource discovery." The process they describe is mostly automatic, i.e., in an electronic environment, the computer can follow specified criteria to form a collection with resources from a large number of sites. Does this mean that the computer can totally replace human intermediaries?

A closer look reveals many limitations of this model. First, it heavily relies on document coding. If the source coding of a document is unfamiliar to the collection program in the computer, it will be excluded regardless of its appropriateness for this collection. Another item may become part of the collection only because, in the computer's view, it possesses the right element—written by the document author, intentionally or mistakenly. Second, the applicability of this model at the current time is questionable. Lagoze and Fielding give the example of a collection of computer science research reports and papers from 120+ institutions that have agreed to participate. There are only three very simple criteria in this collection program: computer science-related, research report, and publishing authority (a computer science department or institute that is a participating member). In addition, all members have to download a program and apply it to managing all documents for this collection. The simplicity and confined scope of this collection is far from the reality of less controllable environments.

Clearly, the computer will slowly take over some aspects of intermediation. However, in collection development, subjective elements in the document, such as the quality of content and the author's viewpoint, are unlikely candidates for automatic processing. Political and personal considerations in an institution, common in collection decisions (Carrigan, 1996; Lee, 1997), are also beyond a computer program. The issues of human versus machine intermediation are worthy of future exploration.

Transient Boundaries in Electronic Environment

The distinction between a collection and a document has become questionable due to the transient nature of electronic resources. Some people perceive a set of retrieved electronic resources in a search session as a new document (Schamber, 1996), and others see it as a temporary collection (Hill et al., 1999). The lack of physical boundaries around electronic objects makes it easier to look at a set of retrieved items as a new unit, either a document or a collection. These variant views prove boundaries in the virtual environment problematic and illusive. Thus, the issue of boundary drawing, for both a document and a collection, needs special attention.

Hypertext technology brings an entirely new horde of issues. It especially tests the boundaries of a document and of a collection. Very commonly, a hypertext document has a number of hyperlinks that allow users to navigate from this document to other related documents. It also may have hyperlinks to the subfiles that its author intentionally made as an intrinsic part of the same document. To the user, the difference between the two may be negligible. For the collection developer, this linking power challenges the integrity of a collection.

Let us assume that Item A has a hyperlink to Item B. The collection developer at one time decides to select A for Collection X. Is B, then, automatically part of Collection X? Even though the collection developer does not explicitly select B, it may be conveniently accessed by users through A. It is unclear whether or not the collection developer has an intention to include B through A. This is not a superficial issue; it embodies a number of professional and legal implications that are of concern. Some collection developers and authors on the WWW have provided a disclaimer to

avoid potential controversy. This practice seems to represent an attempt to draw boundaries for a collection, but realistically, it is more an administrative and legal device than a conceptual definition.

The ambiguous boundaries of electronic documents and collections induce serious complications. Due to the fact that electronic information can be easily manipulated, the entire digital world and its components may manifest themselves as mutable and entangled layers of collections, subcollections, documents, and subdocuments. This instability and ambiguity inherent in electronic resources is affecting every aspect of information work. Information scientists and professionals need a better grip on the issue to succeed in this new and diverse environment.

A Proposal for an Expanded Concept

The forgoing discussion demonstrates the need for an enriched concept of collection in library and information science that incorporates the essential characteristics of the traditional concepts of collection and accommodates the changes brought about by technological advancements as well. Given the new reality created by these advancements, it would be beneficial to broaden the concept of collection to reflect the continuity and interconnectivity characteristic of the information world. The following proposal for such an expanded concept of collection is intended as a first step toward improving collection development, and is meant to be useful for collection developers.

There are two major conceptual frameworks at work here: (1) a view of information seeking as contextual and interactive, and (2) a user-centered approach. First, any collection forms a context that presents to the user a group of selected and organized information resources. The context is sometimes physical, sometimes institutional, and sometimes intellectual. The user interacts with information resources in this context to find relevant information, to learn, and frequently, to explore new ideas. Some users stay within the confinement of this context, but many others reach out to other collections (contexts) as they seek information.

Second, when users are a central concern, their perspective can be directly incorporated. This approach leads to a hypothesis that users and collection developers have differing concepts of a collection. It is hypothesized that the collection developer (human or computer) views a collection in terms of levels of control and the user does so in terms of levels of access.[6] The levels of control may possibly fall into five categories: ownership, lease, interlibrary loan, referral to another collection, and no availability. The levels of access for a user may not match snugly the levels of control and may range from immediate access, to access with a waiting period, to no access at all. Further, it is possible that some users only see two levels: immediate and no access.

The proposed concept of collection below is intentionally inclusive to acknowledge that a collection is not an isolated totality of selected information resources. There are layers of control and layers of accessibility, with built-in interconnectivity. The proposal has two points of view: one from the collection developer, and the other from the user.

Developer

Collections of information resources should be developed with a policy that provides all necessary parameters for the collections, including a definition of the user community, collection scope, format, and depth. Collections have interrelationships with each other, and some collections are subsets of others. The collection developer carefully manages a collection and its subcollections, providing the intended users with immediate and convenient access to information. At the same time, the developer collaborates with other information services in mutual collection sharing, thus broadening collection efforts to assist users in accessing a wider range of resources. Resources selected for the immediate collection may have a variety of formats, and may be physically dispersed and owned by various parties. All resources in the immediate collection should be selected, organized, and made easily accessible through one integrated retrieval system that also supports navigation across collections.

User

Collections facilitate information seeking. An immediate collection provides the first level of access where users of the community may retrieve needed quality information readily and conveniently. This collection must not exist in isolation, and its users should have the flexibility and support to reach beyond this level of access when they need to do so.

The key elements in the above proposal are: a group of information resources, a defined user community, a collection development policy statement, and an integrated retrieval system. The first element, a group of information resources, is the minimum requirement for a collection. The other three are vital for developing collections to satisfy users' information needs. For both users and collection developers, information resources in the collections should have intrinsic quality[7] as well as potential usefulness for the user community (Baker, 1994). Further, functional collec-

[6] This idea is an expansion on Buckland's thinking (1995) regarding degrees of accessibility and privileging. In his article, however, Buckland makes no attempt to scrutinize collections from two perspectives, user's and collection developer's.

[7] The judgment on the quality of information resources often is subjective. A small group of articles about the literary canon and collection development exemplifies this (e.g., Heinzkill, 1990). The issue is very complex and thus is excluded from this discussion. Nevertheless, librarians generally agree that quality must be a consideration in collection development.

tions require a good, comprehensive retrieval system and built-in support to assist users in locating any information resources in the immediate collection and navigating across various levels of access.

A well-designed, integrated information retrieval system (IRS) may have increased importance in the new environment, entirely or partially virtual. First, virtual documents, unlike their traditional counterpart, must have representation in the IRS to be accessible, due to their lack of physical accessibility. This new attribute compels us to rethink the coverage and design of the existing IRS. The most desirable IRS has a comprehensive coverage of all documents in the collection. By including all, not just some, documents, the IRS strengthens the integrity and accessibility of the collection, making the user aware of the full extent and depth of the collection. Second, there is more flexibility in arranging virtual documents or document surrogates in the IRS. Subcollections can be formed not only for subgroups within the user community on a more permanent basis resembling branch libraries, but also temporarily for individual users by following a number of user-identified parameters. Many library OPACs already have similar capacity to allow users to limit searches in a subgroup of catalog records as if they are searching a subcollection. Third, users gain more power in accessing information outside the immediate collection through one conveniently configured IRS, with levels of access clearly indicated.

The above expanded concept of collection makes it unequivocal that tangibility, physical collocation, format, and ownership are no longer adequate for conceptualizing a collection. Unfortunately, they have deep roots in the traditional thinking, and will take some effort on our part to get rid of them in developing and broadening collections. To serve new generations of users, collections must reflect users' choices of information resources, whether printed materials or electronic documents, owned by one organization or distributed on separate computer servers that are thousands of miles apart.

This new concept needs to be tested by future research. Its intended purpose is for the improvement of collection development in particular and information services in general. It also presents challenges to the library operations research mentioned above, and warrants a new edition of library statistics manual that will include reporting materials beyond those that are physically owned by the library. On the other hand, there is never an intention to apply the proposed concept to qualify or disqualify existing collections. Determining whether a collection qualifies, in fact, to be a collection is of little concern. What is more constructive and beneficial is for us to think about developing functional, effective, and accessible collections that users will appreciate and frequent.

Implications for Future Research

Rigorous research is needed to strengthen our understanding of the role and function of a collection and, in turn, to improve collection development. A study of how users view and use collections is a logical first step. As a field, we have accumulated some knowledge about users and their information needs. For example, recent studies have concentrated on information needs and information seeking of specific types of users (Fidel et al., 1999; Thompson, 1997; Westbrook, 1999). What is lacking is an explicit view of the collection as an information-seeking context. This view is of critical importance in forging a research agenda with the following highlighted questions: Does a collection provide a major locus for information seeking? How heavily do users rely on the collection for their information needs? What do users do when the information resource sought is unavailable in the collection? How frequently do they resort to substitute resources in the same collection even when the substitutes are an imperfect match? How does the size and organization of a collection affect information seeking?

Other questions are of particular relevance in the online environment. How do users use electronic collections? Do they prefer navigating freely in the cyberspace by themselves or starting with a collection as they do in a traditional library? If provided with a collection service on-line that is similar to a finite collection provided traditionally, how heavily do they rely on this on-line collection for their information needs? How does the size and organization of a virtual collection affect information seeking? If information is available in both the traditional and the digital formats, which does the user prefer and why? Do preferences depend on the type of information or type of user and in what way?

Research to answer these questions is important for three reasons. First, it acknowledges the importance of a collection as an information *context*, not just as a group of objects. Second, it considers users' perception of collection, in addition to that of collection developers. Third, this research closely ties information seeking to the collection and focuses on how users interact with the collection. Information seeking in this research is rightfully treated as contextual and interactive. Information seekers and collection developers are recognized as capable human beings with varying perspectives and ideas. This line of research will bring new light to collection development by making it a true user-centered process and by adapting it to environments beyond traditional libraries.

Acknowledgment

The author wishes to acknowledge, with many thanks, the critical and constructive comments given by Drs. Raya Fidel, Allyson Carlyle, and the anonymous referees.

References

Alesandrini, K. (1992). Survive information overload: The 7 best ways to manage your workload by seeing the big picture. Homewood, IL: Business One Irwin.

American Association of School Librarians, & Association for Educational Communications and Technology. (1998). Information power: Building partnerships for learning. Chicago: American Library Association.

American Library Association. Reference and Adult Services Division. Collection Development Policies Committee. (1993). The relevance of collection development policies: Definition, necessity, and applications. RQ, 33(1), 65–74.
American Library Association. Statistics Coordinating Project. (1966). Library statistics: A handbook of concepts, definitions, and terminology. Chicago: The Association.
Association of Research Libraries. (1995). Definition and purposes of a digital library [World Wide Web]. Available: http://www.ifla.org/documents/libraries/net/arl-dlib.txt. Accessed Oct. 24, 1999].
Atkinson, R. (1996). Library functions, scholarly communication, and the foundation of the digital library: Laying claim to the control zone. Library Quarterly, 66(3), 239–265.
Baker, S.L. (1994). Quality and demand: The basis for fiction collection assessment. Collection Building, 13(2/3), 65–68.
Buckland, M. (1992). Redesigning library services: A manifesto. Chicago: American Library Association.
Buckland, M. (1995). What will collection developers do? Information Technology and Libraries, 14(3), 155–159.
Carrigan, D.P. (1996). Data-guided collection development: A promise unfulfilled. College & Research Libraries, 57(5), 429–437.
Curley, A., & Broderick, D. (1985). Building library collections (6th ed.). Metuchen, NJ: Scarecrow Press.
Evans, G.E. (1995). Developing library and information center collections (3rd ed.). Englewood, CO: Libraries Unlimited.
Fidel, R., Davies, R.K., Douglass, M.H., Holder, J.K., Hopkins, C.J., Kushner, E.J., Miyagishima, B.K., & Toney, C.D. (1999). A visit to the information mall: Web searching behavior of high school students. Journal of the American Society for Information Science, 50(1), 24–37.
Gardner, R.K. (1981). Library collections: Their origins, selection, and development. New York: McGraw-Hill.
Gorman, M. (1997). Ownership and access: A new idea of "collection." College & Research Libraries News, 58(7), 498–499.
Haines, H.E. (1950). Living with books (2nd ed.). New York: Columbia University Press.
Hardy, A.P. (1982). The selection of channels when seeking information: Cost-benefit vs. least effort. Information Processing and Management, 18(6), 289–293.
Heinzkill, R. (1990). The literary canon and collection building. Collection Building, 13(1/2), 51–64.
Hill, L.L., Janee, G., Dolin, R., Frew, J., & Larsgaard, M. (1999). Collection metadata solutions for digital library applications. Journal of the American Society for Information Science, 50(13), 1169–1181.
International Federation of Library Associations, & UNESCO. (1977). The national bibliography: Present role and future developments. Hague: IFLA. ERIC ED149716.
Kent, A., & Lancour, H. (Eds.). (1971). Encyclopedia of library and information science. New York: Marcel Dekker.
Kuny, T., & Cleveland, G. (1998). The digital library: Myths and challenges. IFLA Journal, 24(2), 107–113.
Kurzweil, R. (1992). The future of libraries, part 3: The virtual library. Library Journal, 117(5), 63–64.
Lagoze, C., & Fielding, D. (November, 1998). Defining collections in distributed digital libraries. D-Lib Magazine. Retrieved January 20, 1999 from the World Wide Web: http://www.dlib.org/dlib/november98/lagoze/11lagoze.html.
Lee, H.-L. (1993). The library space problem, future demand, and collection control. Library Resources & Technical Services, 37(2), 147–166.
Lee, H.-L. (1997). Toward a reconceptualization of collection development: A study of the collecting of women's studies materials by a university library system. Unpublished Ph.D. dissertation, Rutgers, The State University of New Jersey.
Lee, S.H. (Ed.). (1994). Access, ownership and resource sharing. New York: Haworth Press.
Lively, L. (1996). Managing information overload. New York: AMACOM.
Lynch, M.J. (1981). Library data collection handbook. Chicago: American Library Association.
McColvin, L. (1925). The theory of book selection for public libraries. London: Grafton.
Peek, R. (1998). Miss Web Manners on digital libraries. Information Today, 15(7, July/August), 36.
Quinly, W.J. (1956). Audio-visual materials in the library. Library Trends, 5(2), 294–301.
Saffady, W. (1995). Digital library concepts and technologies for the management of library collections: An analysis of methods and costs. Library Technology Reports, 31(3), 221–380.
Schamber, L. (1996). What is a document?: Rethinking the concept in uneasy times. Journal of the American Society for Information Science, 47(9), 669–671.
Scholtz, J.C. (1989). Developing and maintaining video collections in libraries. Santa Barbara, CA: ABC-Clio Press.
Spiller, D. (1986). Book selection: An introduction to principles and practice (4th ed.). London: Clive Bingley.
Taylor, R.S. (1986). Value-added processes in information systems. Norwood, NJ: Ablex.
Thompson, M.L. (1997). Characteristics of information resources preferred by primary care physicians. Bulletin of the Medical Library Association, 85(2), 187–192.
Truesdell, C.B. (1994). Is access a viable alternative to ownership?: A review of access performance. Journal of Academic Librarianship, 20(4), 200–206.
Trueswell, R.W. (1969). User circulation satisfaction vs. size of holdings at three academic libraries. College & Research Libraries, 30(3), 204–213.
Waters, D. (1992). The distributed national collection, conspectus, resource sharing and cooperative collection development. Australian Academic & Research Libraries, 23(1), 20–24.
Waters, D.J. (July/August, 1998). What are digital libraries? CLIR Issues, 4, 1, 5–6.
Westbrook, L. (1999). Interdisciplinary information seeking in women's studies. Jefferson, NC: McFarland.
White, G.W., & Crawford, G.A. (1997). Developing an electronic information resources collection development policy. Collection Building, 16(2), 53–57.
Wilson, P. (1996). Interdisciplinary research and information overload. Library Trends, 45(2), 192–203.