Hybridity, Mutability, Multiplicity: Theorizing Electronic Library Collections

MARLENE MANOFF

ABSTRACT

AN INCREASINGLY COMPLEX ARRAY OF ELECTRONIC OBJECTS is challenging conventional categories and distinctions central to library organization. Extraordinary efforts are being made to integrate these hybrids and new forms into traditional library systems and to recreate the stability of the print environment. This article is an attempt to explore some of the conceptual and theoretical issues raised by the proliferation of electronic objects and to suggest that the current situation calls for a reconceptualization of collection development and bibliographic access.

INTRODUCTION

In the past decade, libraries have undergone a process, not unlike other postmodern institutions—a blurring of categorical distinctions, a melting or erosion of the boundaries between the kinds of work that they do and between the kinds of objects to which they provide access. Online catalogs and library Web pages begin to merge with electronic journals and full-text databases to which they connect. Patrons may find it difficult to distinguish catalogs from indexes, from full-text databases, from document delivery, from interlibrary loan, and even online reference. This discussion is an attempt to grapple with the fascinating conceptual and theoretical questions raised by the proliferation of new electronic objects and to indicate how our understanding of these issues is just developing.

Marlene Manoff, MIT Humanities Library 14S-200, Massachusetts Institute of Technology, Cambridge, MA 02139
LIBRARY TRENDS, Vol. 49, No. 1, Summer 2000, pp. 857-876

Hybridity: The Erosion of Boundaries

We are living in an era of hybrids and provisional boundaries that is remaking the face of academic libraries. Although there has always been a certain ambiguity in the distinctions between collections work, reference, and technical services, technology is accelerating the dissolution of the lines among these functions. Assembling and maintaining a group of electronic products on campus networks, for example, may require the expertise of staff in reference, collections, acquisitions, and systems. For years libraries have been dealing with some degree of overlapping functions, but the demands of the digital environment—electronic reserves, Web resources, virtual collections, electronic journals, and digital preservation—require an integration of functions and expertise sometimes at odds with traditional library structures.

We are also experiencing hybridization on a larger scale. The boundaries between our libraries and the rest of the world are becoming more porous. As more scholarship is being made available through the Internet, we are witnessing a disappearance of the boundaries between academia and the rest of the world. The scholar and his work have become nodes on a vast information space that both integrates and confuses commerce and culture. We are similarly witnessing an erosion of the distinction between the library and the network. Patrons often cannot distinguish precisely where a library's Web site ends and where resources mounted elsewhere begin. The difficulty in determining where online material one is consulting actually resides means that patrons may not know what connection, if any, such material has to the library. The ambiguous boundaries of the Internet make it increasingly difficult for us to map the vast decentered communications network in which we operate.

As library Web sites get slicker and more polished, they begin to look like the corporate and commercial sites to which they are linked. Library networks and Web sites maintain connections to scores of both commercial and academic sites, as well as government, other institutional sites, and even privately maintained pages. These networks of links reveal the embeddedness of contemporary academic libraries within the larger Web of corporate, business, government, and entertainment culture. The fading of the distinctions between these domains, like the fading of the distinctions between elite and popular culture, between news and entertainment, or between art and commerce are hallmarks of the postmodern. Even the language describing students as consumers and library users as customers is evidence of the blurring of corporate and academic worlds.

The Internet is one of the primary places to witness the erosion of the boundaries between advertising and information. As a recent *New York Times* article points out, one may be hard pressed to distinguish advertising from other Web content (Hansell & Harmon, 1999). The *New York Times* reproduces the results of a search on Lycos which demonstrates the challenge of

distinguishing the paid links from the actual search results. Whereas in the print world there are conventions that help in separating ads from editorial content, no such conventions (yet?) exist for online information. What most of us find on the net is not the fulfillment of a utopian fantasy of free information but rather a vast array of consumer choices in a sea of poorly organized information. Not only is it difficult to separate the ads from the content, but it is hard to determine the reliability of information on any site, even with the help of domains ending in .edu, .org, or .com.

Large electronic databases made available by libraries are themselves part of this hybridization, blurring the distinctions among indexes, articles, electronic journals, and other information. Users of these databases can move easily among journals, articles, texts, and material cited in notes. As there is no physical boundary between them, electronic texts are less isolated and independent than those enclosed in discrete physical objects. But it is also less clear to computer users whether they are navigating within a document or between documents or whether they are searching one database or several. They have easier access to more information but it is more difficult for them to distinguish the precise relationship between the pieces of data or the texts they may summon to their computer screens.

Why is the fluidity or instability of electronic text an issue for libraries? For one thing, it raises the question of the extent to which libraries as we know them depend on the existence of discrete objects. Or, to put it another way, our collections and our catalogs—the very structure of traditional libraries—has been based on a system of accumulating clearly demarcated objects. The stacks, reserve rooms, circulation desks, reference areas, and our very buildings were designed to house and circulate books and other physical objects.

As libraries devote a growing proportion of their resources to providing access to large databases, they must address the difficulty of integrating them into traditional library structures. Databases cannot be interfiled with books and journals, assigning them call numbers is not particularly useful, and there is as yet no agreement on whether it is worthwhile or even possible to catalog the thousands of journals, articles, or documents they include.

What are we collecting? Is the object the database or is it the journal, the article, or the other data that it contains? The more we purchase databases of electronic journals or compilations of electronic texts or data, the fewer choices we have about the specific documents or information we acquire. We buy the whole package or none of it. The larger the package, the more likely it is that it will contain a considerable amount of material that we would not otherwise have chosen. It is also more likely that it will duplicate some material that we receive from other sources. These large databases also make it more difficult to provide precise item

level access through current catalog structures. The result is a greatly decreased level of control over what we acquire and how we provide access. Moreover, this lack of control is exacerbated by the fact that many of our electronic subscriptions are leasing arrangements that do not guarantee us permanent possession. Add to this the uncertainty over who, if anyone, will be archiving and preserving much of the electronic material currently available, and the result is a deeply unsettling situation for contemporary academic and research libraries.

Subject specialists building academic and research library collections are feeling increasingly disenfranchised by new modes of acquisition necessitated by giant database packages. As this author has discussed elsewhere, the growth of electronic resources has slowly eroded the traditional model where collections decisions are made by individual subject specialists (Manoff, 1997). Decisions about large expensive electronic databases, especially ones that are multi-disciplinary in nature, typically require input from a number of subject specialists and also staff with technical expertise as well as staff knowledgeable about licensing and copyright. Many libraries have relinquished decision making for large electronic databases to committees representing various kinds of expertise. This broader input may lead to better results, but it also leads to the bureaucratization of the collections process and a dilution of the input of individual subject specialists. Consortial purchases further complicate this system (Manoff, 1997, p. 203).

One way subject specialists assert control over material made available on the Internet in their areas of expertise is through the creation of subject-based Web pages. Although these pages provide a new field in which selectors may exercise their evaluative skills, such work involves new kinds of decision making. If, for example, a history selector identifies a number of sites she deems important enough to link to her history page, she is, in effect, also selecting every site that is linked to by those sites and every site that is linked to by those sites, and so on. She cannot know whether any site she links to will change or disappear, and she has no control over the links that continue to be created or deleted by those sites. Users consulting library Web pages may find it difficult to distinguish remote sites from those created by library staff. Given the looseness of the culture of the net, library patrons are likely to take this in stride despite their confusion. But for selectors, control over the collections that they have built is slowly being drained as these collections bleed out into the vast data spaces of the Internet. Selectors creating subject Web pages are not so much building collections as creating paths out of their collections to resources provided elsewhere. This is a new way to think about collection development.

The Nature of Electronic Objects

Librarians are not the only ones being forced to address the confusion over the nature of the electronic object. George Landow (1992), in

his pioneering book on hypertext and literary theory, devotes a section to the question "What Is the Object We Read, And What Is a Text in Hypertext?" He provides no definitive answer, only an elaboration of the terminological complexities of the online environment and an appreciation of how tenuous is our grasp of textual objects. Landow claims that as "soon as one converts the printed text to an electronic one, it no longer possesses the same kind of textuality" (p. 43). He finds that the tendency to think of these materials as electronic books demonstrates a misunderstanding of the relation between the machine and the text (p. 41).

The difficulty stems, in part, from the immateriality of the electronic word. It seems to exist everywhere and nowhere. And because databases allow users to combine and sort information in many different ways, each search or command may summon or constitute a new object. Confusion also arises from the fact that, when dealing with electronic text, there is no such thing as an original. All instantiations of an electronic text are copies. And the fact that such text can be duplicated, cut and pasted, reassembled and transmitted almost effortlessly online may be unsettling or disorienting. Moreover, electronic text has come to represent a broad range of objects that may now possess color, shape, sound, and movement. Landow claims that, since the electronic medium radically alters the meaning of reading, writing, and text, it becomes increasingly problematic to use those terms when discussing electronic material (p. 41).

As mentioned earlier, libraries are confronting this elusiveness and instability of electronic material through the provision of full-text databases and large packages of online journals. Whether they attempt to provide cataloging or merely pointers in a library Web site to electronic journals or other kinds of documents contained in their databases, they find that the contents of these databases may alter. It is not just a question of deciding which parts of a database are worth identifying for more direct or additional modes of access (although this is not a simple matter either). Database providers like Lexis Nexis and Dialog (Knight Ridder) do not guarantee that their products will remain consistent over time. Items sometimes disappear from these products without warning and therefore libraries may literally be providing access to a different "object" from one day to the next. Is the term "collection development" still appropriate in such an environment? "Content instability" and "content erosion" have become the new buzzwords when dealing with aggregators of electronic journals or other texts. This terminology reflects the anxiety provoked by both the new electronic marketplace and the electronic medium itself. Large databases may be archetypically postmodern in their absolute resistance to containment or control. But, despite the fact that their indeterminacy or lack of fixity defies our attempts to consolidate them within traditional bibliographic structures, libraries persist in their efforts to tame these information monsters.

The problem stems in part from the difficulty of specifying precisely what constitutes a text or a document. Library systems for organizing information are predicated on definitions that will allow us to name and distinguish the objects we provide and decide what constitutes difference. The electronic environment presents us with new kinds of questions involving the boundaries of the electronic object. Is a Web site or a database a single object? What constitutes a new edition of an electronic text, a Web site, or a database? Are the electronic and paper versions of a journal or a text the same object? Is the same Web site viewed through two different browsers the same object even if the site looks different in each? Are Mac and PC versions of the same text the same object? Is the desktop icon for an electronic product part of the object? Are the installation or configuration instructions part of the object? If we are unsure of precisely what constitutes the object, how can we know what we are selecting, acquiring, cataloging, storing, or archiving? What does it mean to perform collection development in an environment of dynamic and volatile objects?

New areas of research are evolving to address the complex issues surrounding electronic text. A whole body of literature is developing to answer the question "what is a document?" As documents are central to law, commerce, education, and government, they play an essential role in maintaining the social order. Moreover, they have been the basis for knowledge management for thousands of years (Renear, 1997a). But, in the electronic environment, we are witnessing the dilution and expansion of the definition of a document to include Web pages, computer files, spread sheets, and so on. As David Levy (1999) argues, this is a source of considerable disorientation and anxiety and not just for librarians and scholars (p. 17).

Researchers are also addressing the related issue of the mutability of digital documents. Recognizing that new information replaces and often eradicates old information, some advocate a focus on the stability of electronic documents at the structural level. They theorize that if they cannot control content, they can at least standardize the document structures that contain content and possibly link these structures to metadata. A number of these researchers have adopted the literary concept of genre and use it as a way to identify and describe documents whose content may change over time. Although this taxonomy is still in its infancy, more intricately developed notions of digital genres may eventually provide some degree of stability for the transmission of electronic text (Yates & Sumner, 1997). This research is a measure of the tenacity with which researchers are seeking to contain the volatility of digital documents.

Another instance of the mutability of our objects is evident in the transition from card catalogs to Online Public Access Catalogs (OPACs). As these now morph into catalogs of consortial holdings and gateways to online databases and full-text resources, they bear less and less resem-

blance to their print predecessors and take on the shape shifting qualities of the many new electronic objects to which they provide an interface and a classificatory structure.

The computer has assisted us in seeing that the physical embodiment of a work is not the work itself. But libraries have provided access to specific physical manifestations. Traditionally a catalog record has described a particular physical object. Even a CD-ROM or a computer tape is a physical object. But when we begin to catalog resources available on the Internet, there is no physical object to describe. And when there is no physical object, one has much less assurance that the object will remain consistent or unchanged over time. If one catalogs a CD-ROM or a computer tape, one can be fairly certain that one's record will remain accurate for the life of the physical object. If one catalogs an electronic product that is accessible over the Web, one has no such certainty.

But all electronic texts, even those on CD-ROMs or computer tapes, are virtual, not physical, objects. Julia Flanders (1997), one of the editors of a large database of early women's writing, has indicated that the lack of a body is a crucial focus for the anxieties and hopes that attach to the digital medium. Flanders claims that we experience the printed book much like a physical body that provides a grounding for the text that it contains. It offers a stable object of reference, and it secures a sense of cultural authority. With electronic text we experience a "loss of that body [that] can seem like the severing of the bonds between meaning and its foundation; the opening up of the doors of chaos" (pp. 127-28). This may be part of the compulsion of many people to print out and save documents that will presumably remain available electronically. And it may be why faculty are so frequently opposed to discarding or storing material available online. They want the security of knowing that they can consult the paper copies even if they have no artefactual value. They still want to be able to hold the text in their hands.

Or perhaps it is also the case, as Stuart Moulthrop (1995) proposes, that scholars are resisting what he calls "the threat of multiplicity in electronic writing" (p. 58). Like many other theorists, Moulthrop finds that writers have always struggled against the linearity of print. Digital technology, and more particularly hypertext, can be seen as the culmination of that struggle. Hypertext, with its surfeit of narrative possibilities, becomes the fulfillment of the postmodern dream of multiplicity. It signals an end to linear narrative as it consistently seduces the reader into clicking on yet another link and interrupting the narrative flow. Reading thus becomes a process of continually detouring and refocusing as each new link introduces a new path or option. Reading in such an environment becomes a different kind of activity. Moulthrop theorizes that the growth of electronic networks will in its turn inspire a resistance to the polyvalence of hypertext and a desire for a return to the linearity of print (p. 58).

Moulthrop (1995) sees the multiplicity of hypertext as especially problematic for literary criticism because, as he puts it, in "its root sense, 'criticism' implies a separation of one discourse from another" (p. 59). In hypertext environments, the whole point is to be able to move seamlessly between documents. Primary and secondary sources are much harder to distinguish as author and critic occupy the same textual space. But the integration of the text and its commentaries made possible by hypertext has implications for many fields besides literature. Large electronic archives have a kind of leveling effect in that they not only erase distinctions between primary and secondary texts but also between marginal and mainstream or canonical and non-canonical. Moulthrop finds that if "one chooses to work in hypertext, one has no clear defense against the potential vastness of the network and its multiplicity, if not 'randomness'" (p. 59). Whereas the linearity of print may constrain, the plurality of electronic text may overwhelm and confuse. Moulthrop insists we will struggle against the Web as surely as we have struggled against the line.

Accessing and Organizing Electronic Texts

For librarians and scholars, the anxiety attached to the mutability and multiplicity of electronic text is exacerbated by the sheer volume of electronic information. As Kathryn Sutherland (1997) points out, "on the one side, there is the information revolution, anarchic, global, culturally leveling, and largely uncritical in its methods; on the other, there is scholarship, selective, judgmental, and exclusive in its cultural priorities" (p. 11). The lack of selectivity or standardization on the Internet and the difficulty in discerning the authority or reliability of material located through browsers or search engines contributes to the sense of both scholars and librarians that the electronic revolution may be antithetical to serious scholarship. Some faculty members see the Internet as desperately in need of professional or scholarly editing, much as librarians see it as desperately in need of bibliographic ordering. The situation may seem to invite extraordinary measures to assert control over this deluge of information.

Much of our current energy in libraries is directed at trying to establish in the electronic environment the kind of control we had in the paper one. We are cataloging electronic resources and providing links to this material in our OPACs. We are creating our own Web pages where we attempt to organize material by subject or provide digital versions of local material. A number of libraries are participating in the CORC (Cooperative Online Resource Catalog) project to experiment with cataloging material available on the World Wide Web. But there is also tremendous pressure on libraries to rethink their assumptions about control. Projects like CORC may enable libraries to catalog Web resources, but they will nevertheless be providing catalog records for material that exists on remote

Internet sites which they do not manage or direct and which may change or cease to exist.

CORC brochures describe the project as applying "the traditional practices and principles of librarianship" to Web materials, including authority control. But even if CORC can provide high quality cataloging (and that remains to be seen), is authority control or even bibliographic control as we have known it a viable option? When applied to material on the Internet, isn't it a lot like trying to nail Jello to a wall? Anyone who has written an article or book for print publication in the last few years and who has included citations to online information may be familiar with the experience of having a Web site or two disappear or change before their reference to it has even made it into print. And how does one cite postings to news groups and chat groups when the posters may call themselves bizarre names and when the postings appear fleeting enough to be gone long before one's readers might think to search for them? And how does one cite an electronic source when the new style guides are obsolete before they make it into print? No matter how much libraries refine and improve procedures for cataloging Web resources, the mutability of those resources will prevent the achievement of anything comparable to the level of bibliographic control possible for printed objects.

James O'Donnell (1998), a classical scholar as well as a theorist of hypertext, declares that over time we will lose the sense that "discourse must be fixed to be valid" (p. 41). O'Donnell contends that we will adjust to the fluidity of our electronic objects and will learn to celebrate it. He tells us that there is hardly anything he has published in his fifteen or so years as an academic that he would not change if he could: "[W]ords that I know to be inadequate and in some cases untrue continue to speak for me. I am no longer the person I was when I wrote them, but I am still somehow their author" (p. 41). The print medium demands and ensures a level of fixity in scholarly and popular discourse that is at odds with contemporary notions of the fluidity and mutability of the subject. It is relatively easy to alter one's electronic publications as one's views or one's knowledge grows or changes. Web sites can be updated and electronic text is easily edited. But one's print publications will continue to proclaim one's no longer current views as long as they remain in circulation.

Many find that this is exactly what is valuable about the printed word. It would seem to be the ideal medium to document the historical record as it is less open to erasure or change than the electronic word. But interestingly enough, in some instances, the electronic medium provides a superior means of representing and documenting the evolution of a text. Digitization has made a tremendous contribution to the creation of scholarly editions of literary works precisely because it allows for the inclusion of multiple versions of a work (including facsimile reproductions) and because with split screens one can easily compare versions. Creating print

editions of Wordsworth's poems, for example, has been complicated by the fact that the author continued to revise well into old age poems he had written in his twenties. The multiple versions of *Lyrical Ballads* (http://www.dal.ca/~etc/lballads/) seem to be ideally suited to an electronic representation as they allow one to view and compare variant editions. Thus, although the electronic medium may threaten our ability to maintain a complete historical record, it does enhance our ability to more fully represent and disseminate portions of that record.

Addressing the issue of the multiplicity and lack of fixity of the electronic word, but from a library perspective, Ross Atkinson (1998) sees the problems it creates for bibliographic access. Although he acknowledges the benefits of the electronic environment, he indicates that digital culture "is rather more habituated to, and accepting of, some loss of intellectual content. The digital culture is characterized by information extracted from remote sites, of which the local user has little knowledge and even less control; it is a culture of Web sites that change every day without warning" (p. 10). Although he finds this tolerance for intellectual loss to be in conflict with traditional library culture, Atkinson also declares that libraries will inevitably be forced to adapt: "Some loss . . . is nevertheless becoming increasingly understood as part of the price of digital access" (p. 10). So, although libraries are resisting, it has become a simple fact of life that they provide access to a body of electronic information in constant flux. Bibliographic control takes on a whole new meaning in this environment. Does it become something different when it is aimed at moving targets or constantly metamorphosing objects? Is it still collection development and cataloging or is it some new hybrid? Perhaps we need to ask how much the theories and practices of collection development and cataloging can be stretched to accommodate the electronic environment. At what point might we need to think about new theories as well as new practices?

In yet other ways digitization has complicated our understanding of textual objects. Scholars involved in digitizing primary sources in the humanities have repeatedly noted that converting a text to electronic form typically alters one's view of the original materials. In order to make these texts functional—i.e., searchable or manipulable—their distinct features are tagged or encoded so the computer can distinguish their various elements such as titles, paragraphs, quotations, footnotes, patterns of imagery, and meter or rhyme schemes. Encoding a text essentially involves a cataloging or classification of its parts. Thus, experience with textual encoding has implications for library classification. The most widely used system for humanities text encoding is SGML (Standardized Generalized Markup Language) which emerged as a result of efforts in the early 1980s to develop standardization. SGML allows one to encode texts in such a way as to enable quite sophisticated searching. If, for example, one were tagging a collection of historical documents, tagging might enable one to

search for Washington as city as distinct from Washington the person and even Washington as author (Chestnutt, 1997). Or it could allow one to tag and thus search references to class or gender.

But, as Steven Johnson (1995) has explained, different markup languages presuppose different definitions of what constitutes a literary work (p. 30). One cannot tag everything. What one does tag defines the nature of the text one is presumably only describing. Whoever makes the decisions about how a text will be encoded has power over the meaning of that text. Cataloging and indexing have similar effects in that they help determine the conditions under which a researcher will be able to retrieve a particular text or piece of information.

Having worked extensively with SGML, Allen Renear (1997b) tells us that the structures identified in encoded texts are as much a reflection of the interests of researchers as they are the reflection of anything inherent in the texts themselves (p. 122). Those doing the encoding therefore run the risk of interpretive bias. The choices encoders make about which features to tag will determine what future researchers will be able to do with a text that has been encoded. In effect, features that are not tagged disappear since they cannot be retrieved through searching.

The realization of the subjective nature of the encoding process has led to debates about whether texts have any independent existence at all. Renear (1997b) identifies himself as one of those who retains a belief that texts do possess objective structures independent of the processes of naming and encoding them. But he also describes a group of theorists who have come to share the post-structuralist view that texts have no objective independent existence but are instead the product of the theories and methods that are used to "transcribe, edit, analyze, or encode them" (p. 122). He calls this view Antirealism (p. 121). These theorists believe that it is impossible to create an objective or transparent map of a text—one that merely describes and does not impose a structure. Librarians may well wonder what it means to apply the precision tools of bibliographic control to entities whose objective existence is increasingly being called into question. While Renear does not consider himself to be an antirealist, he acknowledges the instability of textual objects, the coercive power of the systems we use to represent them, and the subjective nature of the encoding process. Perhaps we should be thinking more about the philosophical implications of such theories for library classification.

There has been considerable discussion about how systems like the Library of Congress Classification system, the Library of Congress Subject Headings, and the Dewey Decimal Classification are prone to a kind of coercion not unlike that described by Renear. Confidence in these systems is being undermined by challenges made to the assumption that the universe of knowledge is transparent and can be objectively mapped. There is no longer a widely sustained belief that there is a single nature and

order of things, logically organized and structured as a hierarchy, ready to be embodied in a single classification scheme. Although both LC and DDC (at least implicitly) claim to be extrapolating a preexisting order in the world of knowledge, merely naming what already exists, they are both socially and historically conditioned maps that always leave some territory unnamed and uncharted. Catalogers are left to deal with the gaps and limitations of these systems.

Francis L. Miksa (1998) acknowledges that the implications of the "epistemological characteristic of a post-modern age are especially profound for the classification of knowledge" because such classification "is in its own right an elaborate assertion about the collective knowledge of humankind" (p. 86). Nevertheless Miksa tries to describe a future for the Dewey Decimal Classification system. He argues that since all experience is open to multiple, if not infinite, interpretation, classification systems should be redesigned to more flexibly accommodate a multitude of perspectives (p. 87). Miksa proposes a scheme to enable DDC to do just that. He envisions "a vast array of moveable or interchangeable facets of categories, a system that is perhaps best called an object relational database management system of categories" (p. 89). My own view is that making LCC or DDC more malleable would be considerably more difficult than Miksa imagines, and that it would still not fully address the problems he describes.

Postmodern theory suggests the value of local systems created to meet local needs—i.e., of tools that can be refashioned when no longer useful. Thus one could argue that many different classification schemes will be required to meet many different demands. Such thinking dovetails with arguments for allowing users to create customized interfaces. The current buzz about MYLibrary software (http://my.lib.ncsu.edu/), developed by Eric Lease Morgan at North Carolina State University, suggests that libraries are moving in the direction of acknowledging that their constituencies are diverse enough to require multiple options. In any case, with the growth of large-scale information systems, there has been an explosion in the number of global information schemes. Leigh Star (1996), comparing a number of such schemes, points out how all of these systems struggle with uncertainty, ambiguity, and standardization (p. 5). No one system can resolve all ambiguity and, the broader the system, the greater the likelihood of ambiguity. The advantages of some of these systems include the level of granularity and the specific subject terminology they offer particular fields and projects. The growth in the number of alternatives should encourage librarians to question the value of a one size fits all classification scheme.

In the case of LC, a consistent problem has been the length of time it takes for new subjects and new terminology to be recognized and incorporated. The slowness of LC to adopt new language has to do with what

Earl Lee (1998) calls its "obsession with formal rules and procedures" (p. 97). Lee finds that librarianship resembles the legal profession in that both fields are "based on formalized rules created and approved by authorities" (p. 97). Libraries depend on AACR2 and past cataloging practice in much the same way the legal profession depends on law and precedent (p. 97). Both professions are slow to accommodate new fields and developments. Librarians are still trying to find ways to subsume electronic material within traditional print-based cataloging practices (albeit with some tweaking of AACR2), much as lawyers are still trying to adapt old copyright and intellectual property law to new electronic objects.

Gary Radford (1998) has described what he also sees as the problem of library systems excessively devoted to order and control. He declares that it is no coincidence that librarians are stereotyped as enforcers of silence and policers of order (p. 618). The problem, he says, is that librarians tend to posit a user who "must engage with the rationality of the library directly and must submit to its version of the order of things before" finding what he or she needs (p. 620). Radford contends that librarians mistakenly assume that library research is a purely rational and scientific enterprise. He argues for the importance of imagination in the research process. It is imagination or intuition that enables one to predict which paths will prove most fruitful and how seemingly unrelated materials one retrieves may be assembled to make something new. Seen in these terms, the "librarian's role becomes that of a guide, not only to the preexisting order of the library that comprises its catalogs and indexes but to the creation of new orders developed and made possible by the capabilities of computer searching" (p. 630). Or, to put it another way, library research should be less about forcing people to identify the correct LC-approved subject heading or the correct controlled vocabulary and more about providing the opportunity to forge new connections and find multiple paths through relevant data and information.

Radford's argument echoes one made by Jean-Francois Lyotard (1979/1984) in the late 1970s. Lyotard was one of the first to recognize the significance of the computer for the transmission and creation of new knowledge, but he also formulated an early version of what we have come to call computer literacy. He claimed that the most important thing we will teach future students is "how to use the terminals. On the one hand, that means teaching new languages and on the other, a more refined ability to handle the language game of interrogation—where should the question be addressed, in other words, what is the relevant memory bank for what needs to be known? How should the question be formulated to avoid misunderstandings?" (pp. 50-51).

Lyotard (1979/1984) declares that successful research involves the ability to connect data or information—to devise new arrangements. He finds that this "capacity to articulate what used to be separate can be called

imagination" (p. 52). This conception of research as essentially a creative enterprise, like Radford's (1998), cited earlier, might encourage us to rethink the nature of the maps that we construct for our users and the assumptions we make about how they will proceed.

Since Lyotard, there has been growing recognition that a new kind of information literacy is required. The sheer volume of material available in electronic form requires that we think more about retrieval, navigation, manipulation, and management of information. But what authors like Lyotard and many others outside of the library and information professions don't seem to recognize is that skillful researchers with powerful computers are only part of the equation. Equally important is the creation of information spaces that are easily navigable and conducive to research. Internet search engines, for example, typically yield huge quantities of irrelevant data. Even the most computer or information literate may suffer from inadequate tools.

How to achieve the most effective searchability remains open to debate. If hierarchically organized systems with controlled vocabularies are too rigid and if full-text searching yields masses of inappropriate or irrelevant hits, what are the alternatives? Ideally, tools will be developed that will allow for open-ended searching where the user constructs his or her own path through the maze of information. This means that he or she will not have to follow paths laid out in advance, for example, by LC Classification. The case for such an alternative is eloquently made by Greg Anderson (1992) who claims that the

> goal of libraries and technology is freedom; to enable the reader or author to frame knowledge without constraints and focus energy toward the creation of knowledge rather than on understanding an imposed external organization of that knowledge. Freedom exists when the author/reader can build upon the linkages and paths of knowledge in a flexible, multi-faceted world. (p. 114)

Precisely how to accomplish this is not yet clear although many of the authors cited earlier would seem to be arguing for its creation. We have yet to see new systems that will allow for fine grained searching (like that provided for by LCC and by text encoding schemes like SGML) that do not also force users to accept rigid organizational structures and controlled vocabularies. Is it really possible? Anderson proposes that multimedia information technologies may help libraries to more flexibly represent knowledge through a kind of stratification or layering process that would be more visually effective than current structures (p. 113). But he also acknowledges that the technology does not yet exist to provide the multimedia support he envisions. Like many others who predict new multidimensional information spaces that will enable users to easily grasp relations between large numbers of documents, Anderson does not provide a concrete or detailed description. His most specific suggestion seems to be

that the primary role of the library should be "capturing, creating, and nurturing the linkages, pathways, and management of nodes of information" (p. 115).

Anderson's argument sounds rather like George Landow's (1992) that "we must abandon conceptual systems founded upon ideas of center, margin, hierarchy, and linearity and replace them with ones of multilinearity, nodes, links, and networks" (p. 2). Steven Johnson (1997) offers a similar formulation. He describes the Web as very like the Memex envisioned by Vannevar Bush in the 1940s. The Memex was a major early theorization of an information processor that is now viewed as a precursor to the personal computer. Johnson claims that, to Bush, what made a particular piece of information valuable "was not the overarching class or species that it belonged to but rather the *connections* it had to other data. The Memex wouldn't see the world as a librarian does, as an endless series of items to be filed away on the proper shelf. It would see the world as a *poet* does: a world teeming with associations, minglings, continuities" (p. 119).

This emphasis on connection, as opposed to classification, is also made by Stuart Moulthrop (1993). Observing that all distributed computer systems are hypertextual in nature, Moulthrop elaborates on the effects of navigating a system based on connections, affiliations, and linking (p. 71). He finds that the experience is nicely illustrated by Thomas Pynchon when he defines paranoia as the realization that everything is connected (p. 81, citing Pynchon, *Gravity's Rainbow*, 1973, p. 820). Moulthrop finds that "in dealing with vast and nebulous information networks . . . a certain 'creative paranoia' may be a definite asset" (pp. 82-83). Moulthrop proclaims "the promiscuity of hypertext," noting that the word "promiscuity" comes from a root meaning a tendency to seek relations: hypertext is about "promiscuous, pervasive, and polymorphously perverse connection" (p. 84). Paranoia is valuable in the sense that it is a heightened sensitivity to connections—to recognizing patterns which others might find meaningless. When navigating complex systems of links such as those on the Web, creative paranoia, like intuition and imagination, may be effective in ferreting out, piecing together, and ultimately transforming information into knowledge.

All of these descriptions suggest the limits of elaborately ordered systems like LCC and DDC and point to the possibilities of the electronic environment. Moulthrop (1993) believes that the very multiplicity and promiscuity of massive hypertext systems militate against hierarchy and he sees more advanced systems pointing toward a "'post-hierarchical' information order" (pp. 91, 93). Libraries have, in fact, already begun to provide alternatives. Electronic catalogs allow much greater flexibility in searching than paper catalogs. Many OPACs allow one to locate the book one wants even if one only knows the author's first name and the third

word of the title. Try that with a paper catalog. With sufficient searching options and the ability to combine and sort in a variety of ways, electronic catalogs mitigate, although they do not eliminate, the limits of LC subject hierarchies. Nevertheless, libraries need to do more to foster the development of new and multiple systems of order.

Central to that development will be the creation of new systems of linkage. Steven Johnson (1997) interestingly proposes that the hypertext link is the most important form of punctuation to emerge in centuries (pp. 110-11). He traces the idea of the link to Vannevar Bush (p. 116) and argues that it could provide the basis for a whole new grammar and syntax of connection (p. 111). Its importance for librarians may be the possibilities it offers for establishing relations among ideas, texts, documents, data, and images without imposing a hierarchical structure. Tim Berners-Lee (1989), the creator of the World Wide Web, designed HTML as a system of links to allow for the establishment of multiple connections. He decided against a hierarchical structure precisely because it would foreclose and constrain the emergence of new connections.

Conclusion

Libraries should perhaps regard the confusions and destabilization generated by digital technology as an opportunity to finally address some of the limitations of traditional library access and organization. In the first large-scale study of its kind, researchers recently found depressingly low levels of user understanding of Library of Congress Subject Headings. Users had an overall accuracy rate of only about 36 percent (Drabenstott, Simcox, & Fenton, 1999, p. 158). The authors conclude that drastic change may be necessary and recommend involving library users in the development of new subject headings (p. 159). But perhaps what the study really demonstrates is the problematic nature of controlled vocabularies and the advisability of providing multiple points of access. This is much easier to do in the electronic environment.

There are, for example, interesting possibilities presented by the development of the Dublin Core as well as its adoption by the CORC project. Dublin Core is a metadata set established to provide simplified cataloging of electronic material. It allows for the possibility of using or not using LC Subject Headings or other standardized vocabularies. This could mean tremendous flexibility in classification of electronic documents. Since LCSH was developed primarily to classify monographs and much of the material on the Internet is not monographic, it could be extremely useful to adopt subject thesauri from specific disciplines. But Dublin Core and CORC also present the possibility of developing entirely new ways to describe electronic material since they do not require (although they do recommend) the use of standardized vocabularies. Of course, there are always risks associated with allowing for deviations. However, given that

Dublin Core is being used to catalog electronic objects rather than traditional printed volumes, the need for new options seems obvious. These might well provide an opportunity for some pushing of the envelope on subject headings.

Libraries also have the opportunity to build alternative systems of access. One way for libraries to provide additional orders and modes of connection is for subject specialists to create Web pages in their areas of expertise. These provide an opportunity for selectors to shape new information spaces that are not dependent on LC or standard classification hierarchies. Selectors can construct Web pages that do what LC cannot do. They might even team up with catalogers to do this kind of work. Such a page, for example, could organize material by methodological approach. It could identify resources that represent Freudian, feminist, new historicist, or postcolonial perspectives. Similarly, one could create a page on a controversial topic like abortion and organize resources into pro and con. LC does not make these kinds of distinctions. One cannot determine from LC subject headings whether an author has written from a left wing, right wing, center, or lunatic fringe perspective, but a Web site could adopt such a framework or devise new ones. Subject pages on library Web sites provide ideal spaces to experiment with alternative forms of organization. Not much work has yet been done in this area.

Subject specific Web pages also provide a way of compensating for the weakness of traditional classification schemes in dealing with interdisciplinary areas. A Web page for American studies can pull together resources in history, political science, literature, religion, and sociology. LC and DDC do not effectively bring these resources together because they are built on discipline-based hierarchies that get in the way. Pages offered by subject specialists may be tailored to meet the information needs of local researchers. In institutions with large Asian, African, or Latin American studies programs, for example, such pages help compensate for the difficulty of using LC to navigate these fields. Web pages are also useful for subjects like media studies, cultural studies, Victorian studies (or any other period-based field) that fares poorly with either LC classification or subject headings. Unlike access through OPACs, Web access allows subject specialists to describe strengths and weaknesses of various resources and offer hints about how best to search them. Web pages also provide the opportunity to describe why a particular resource may be useful and to provide evaluative judgments and remarks about authoritativeness. Of course, this kind of work requires that libraries continue to cultivate staff who are knowledgeable about the scholarship they support. And as that scholarship becomes increasingly interdisciplinary, subject specialists will need to broaden their expertise to encompass material in adjacent fields.

Web pages built by selectors may also help compensate for the difficulties of traditional library subject classification by providing discipline

specific searching tips. A page entitled "Literary Theory" on the Bobst Library Web site at NYU, for example, points out that searching the term "literary theory" as a subject in the library catalog will yield nothing (http://www.nyu.edu/library/bobst/research/guides/rg18.htm). The Web site provides a sizable list of relevant subject headings, none of which are simple or intuitive. Not many users would guess that the primary LC subject heading is "Literature—History and Criticism—Theory." Having the help of such a Web page may be the closest one is likely to get to simplified LC access to literary theory.

One might object that library catalogs and Web pages may provide inconsistent descriptions and forms of access for the same resource. Robin Wendler (1999) finds it problematic that Harvard's online catalog and its HOLLIS Plus menu system have completely different subject headings and titles for Lexis Nexis (p. 48). She proposes, as a remedy to such inconsistencies, working to expand the catalog record to create automatically generated Web menus (p. 49). For purposes of updating, there is certainly an advantage in maintaining only one link to each resource and consistency of labeling makes good sense. However, it is not an advantage to prevent selectors from creating customized descriptions and annotations. Selectors should be able, for example, to describe the same resource differently on a literature and a history page so as to advise each constituency about the most appropriate use of that resource. A system that maintains a master set of links and still allows for different resource descriptions and annotations would be a better solution. Wendler repeatedly touts the advantages of coherence and consistency, but I would suggest that one of the major strengths of hypertext systems such as electronic databases and the World Wide Web is the opportunity to provide alternative descriptions and paths to the same place. This is where Web technology can improve on traditional catalog access.

Expanded subject Web pages are just one of the possibilities. As libraries evolve amidst rapid technological change, we need to bear in mind that the best path is not necessarily trying to recreate in the electronic environment a system that mimics as closely as possible the bibliographic modes of access of the print environment. Thus, although it is certainly an honorable goal to bring order to the net, libraries must consider whether providing some modified version of LC or DDC classification to vast numbers of electronic objects is the best way to go about this. Before trying to rein in its uncontrolled vocabulary and seeking to subdue or colonize electronic space armed with a version of the LC classification tables, it might be well to consider what other kinds of contributions we might make to the future of research in the digital environment. Rather than viewing the ambiguity of the electronic object as disabling, we should view it as an opportunity to rethink and reformulate library collections and access.

REFERENCES

Anderson, G. (1992). Dimensions, context, and freedom: The library in the social creation of knowledge. In E. Barrett (Ed.), *Sociomedia: Multimedia, hypermedia, and the social construction of knowledge.* Cambridge: MIT Press.

Atkinson, R. (1998). Managing traditional materials in an online environment: Some definitions and distinctions for a future collection management. *Library Resources & Technical Services, 42*(1), 7-20.

Berners-Lee, T. (1989). *Information management: A proposal* [for the World Wide Web consortium]. Retrieved February 4, 2000 from the World Wide Web: http://www.w3.org/History/1989/proposal-msw.html.

Chestnutt, D. R. (1997). The Model Editions Partnership: "Smart text" and beyond. *D-Lib Magazine,* July/August. Retrieved February 4, 2000 from the World Wide Web: http://www.dlib.org/dlib/july97/07/chestnutt.html.

Drabenstott, K. M.; Simcox, S.; & Fenton, E. G. (1999). End-user understanding of subject headings in library catalogs. *Library Resources & Technical Services, 43*(3), 140-160.

Flanders, J. (1997). The body encoded: Questions of gender and the electronic text. In K. Sutherland (Ed.), *Electronic text: Investigations in method and theory* (pp. 127-143). New York: Oxford University Press.

Hansell, S., & Harmon, A. (1999). Caveat emptor on the Web: Ad and editorial lines blur. *New York Times* [National ed.], February 26, pp. A1, A12.

Johnson, S. (1995). Repossession: An electronic romance. *Lingua Franca, 5*(4), 24-33.

Johnson, S. (1997). *Interface culture: How new technology transforms the way we create and communicate.* [San Francisco]: HarperEdge.

Landow, G. (1992). *Hypertext: The convergence of contemporary critical theory and technology.* Baltimore: Johns Hopkins University Press.

Lee, E. (1998). *Libraries in the age of mediocrity.* Jefferson, NC: McFarland & Company, Inc.

Levy, D. M. (1999). The universe is expanding: Reflections on the social (and cosmic) significance of documents in a digital age. *Bulletin of the American Society for Information Science, 25*(4), 17-20.

Lyotard, J. (1984). *The postmodern condition: A report on knowledge* (G. Bennington & B. Massumi, Trans.). Minneapolis: University of Minnesota Press (original work published in 1979).

Manoff, M. (1997). Cyberhope or cyberhype? Computers and scholarly research. *Canadian Journal of Communications, 22*(3/4), 197-212.

Miksa, F. L. (1998). *The DDC, the universe of knowledge, and the post-modern library.* Albany, NY: Forest Press.

Moulthrop, S. (1993). You say you want a revolution: Hypertext and the laws of media. In E. Amiran & J. Unsworth (Eds.), *Essays in postmodern culture* (pp. 71-97). New York: Oxford University Press.

Moulthrop, S. (1995). Traveling in the breakdown lane: A principle of resistance for hypertext. *Mosaic: A Journal for the Interdisciplinary Study of Literature, 28*(4), 55-77.

O'Donnell, J. J. (1998). *Avatars of the word: From papyrus to cyberspace.* Cambridge, MA: Harvard University Press.

Radford, G. P. (1998). Flaubert, Foucault, and the bibliotheque fantastique: Toward a postmodern epistemology for library science. *Library Trends, 46*(4), 616-634.

Renear, A. (1997a). The digital library research agenda: What's missing—and how humanities textbase projects can help. *D-Lib Magazine,* July/August. Retrieved February 4, 2000 from the World Wide Web: http://www.dlib.org/dlib/july97/07renear.html.

Renear, A. (1997b). Out of praxis: Three (meta)theories of textuality. In K. Sutherland (Ed.), *Electronic text: Investigations in method and theory* (pp. 107-126). New York: Oxford University Press.

Star, S. L. (1996). *To classify is human* [keynote talk presented at Hypertext '96]. Retrieved February 4, 2000 from the World Wide Web: http://alexia.lis.uiuc.edu/~star/hypertext96.talk.final.html.

Sutherland, K. (1996). Looking and knowing: Textual encounters of a postponed kind. In W. Cherniak, M. Deegan, & A. Gibson (Eds.), *Beyond the book: Theory, culture, and the politics of cyberspace* (pp. 11-22). Oxford: Office for Humanities Communication Publications.

Wendler, R. (1999). Branching out: Cataloging skills and functions in the digital age. *Journal of Internet Cataloging, 2*(1), 43-54.

Yates, S. J., & Sumner, T. R. (1997). Digital genres and the new burden of fixity. In R. H. Sprague, Jr. (Ed.), *Proceedings of the Thirtieth Hawaii International Conference on System Sciences* (Volume VI, Digital Documents Track) (pp. 3-12). Los Alamitos, CA: IEEE Computer Society Press.