New Roles for Special Collections on the Network

Peter S. Graham

There will be no special collections on the network in the traditional sense because electronic information is not maintained in artifacts. Special collections have existed to support preservation of the human record as instanced in original documents or in specific documents of importance. Electronic documents, however, do not depend on their physical medium for their importance, nor does their medium provide evidence that assists in better understanding their texts. Special collections will continue in importance because of the continuing importance of artifactual documents. Special collections librarians may have new and distinctive roles in the electronic environment, particularly with respect to intellectual property and in the merging of special and general digital collections.

To begin with, the simple and obvious truth is that there will be no special collections on the network because artifacts will not exist on the network. Special collections, after all, are collections of artifacts, whereas the network and its nodes are repositories of volatile electronic information. However, special collections of books and manuscripts still will be represented on the net. In the first place, there will be surrogates of their holdings: Digitizing projects now are an activity of many rare book and manuscript collections, large and small; and, more important, the skills of rare book librarians and curators will be as essential on the network as they always have been. In fact, the network offers an opportunity to make more evident for the main collections the substantial intellectual contribution of special collections librarians.

Artifacts (once called books[1]) come under the classical definition of mass because they have weight and occupy space; information on the net does not. As artifacts, objects in special collections are of interest because they either contain information inscribed on them or present information in themselves (whether associatively, through provenance, or inherently, perhaps in the way a handpress book has been printed and assembled). Artifacts contain information, including texts and graphics (once called pictures). On the network, users (once called readers) directly manipulate information. The foundation of a collection is information, whether artifactual or electronic. The way a collection is managed, and even defined, depends on the difference.

The Value of Books As Artifacts
Artifacts in libraries can have any of at least three values: They provide informa-

Peter S. Graham is Associate University Librarian at Rutgers University Libraries; e-mail: psgraham@rci.rutgers.edu.

tion about their texts; they provide information about publishing, reading, ownership, and all that we call the history of the book; or they are appealing esthetically.

The esthetic quality is easiest to describe and for present purposes may be accounted for most easily. Occasionally, a book is noted for its esthetic quality, though most often the appearance of books, particularly older ones, is routine if not unsightly to anyone not predisposed to appreciate them. The range of esthetic values can be considerable. Bindings always have had appeal, from early books studded with precious stones to

---

The foundation of a collection is information, whether artifactual or electronic.

---

contemporary hand-tooled leather. Some books are a pleasure to hold and read, such as Aldine octavos, Modern Library and postwar Knopf books, or books from the Spiral Press. Others are intended to offer visual pleasure but are in fact difficult to handle as books, such as oblongs designed by contrarians and most coffee-table books, which are too heavy to hold comfortably. The books most celebrated as beautiful are typically those that add extraordinary design to important textual content, such as the Book of Kells, the Arion *Moby-Dick,* and Bruce Rogers's lectern Bible. In all these cases, it can be seen that there is no obvious analogy to networked texts. Although workstation screens are becoming more sophisticated, they still cannot present information as attractively as a finely made book. No amount of elegance in the construction of laptops or communication protocols will carry weight with those attending to the esthetic experience of reading.

But the primary value of books as artifacts is the way in which they provide information about their texts, both explicitly and implicitly. Readers come to the library to find information in books and,

less often (though importantly), to study the book as object to gain further information about the text or about history more broadly considered. Thus, they use books in their most important artifactual function: as artifacts, whether tablets, manuscripts, or printed books, the materiality transmits the text.

Books embody text. In the interval between the oral tradition and the electronic age, the artifact is the medium of textual transmission. Inevitably, the medium also forms our sense of how the text is transmitted.[2] Roger Chartier, D. F. McKenzie, and Robert Darnton have taught us how size, shape, type face, and portability affect our understanding of the text.[3] Thus, the artifacts, beyond the texts, provide information on the history of the book as a socially created and socially consumed object. This comprises evidence about publishing, printing, binding, and other bookmaking crafts and economic activities, and thus about both authorship and the social organization of work. The physical design and construction of the book provides evidence as to the way readers were intended to perceive it. Often, too, individual books themselves will have a history and become evidence of the transmission of ideas: This book was made in this place at this time, has this reader's name in it, has these marks in it, and was found in this place at this later time.[4]

The use of artifactual evidence for the history of the book, except for practitioners and enthusiasts, is of relatively little interest to most readers. Yet, often the structural and physical evidence that an object supplies is essential for full textual understanding: The typography, page and sheet layout, collation, binding, and all the other aspects of printing history that can be inferred from the object itself all have informed our understanding of the text it carries. The Anglo-American bibliographic tradition has long demonstrated successes in showing how the structure of the book can inform our

understanding of the text itself.[5] The history of the book as a study of social and intellectual history has been similarly informed by examination of typesetting and the sequence of printing, paper construction, paper sources compared to location of printers, format of book in relation to text, and the like.

> There is, therefore, nothing "special" about an electronic collection. There are no artifacts to provide added value to the substantive information.

The artifactual evidence has been emphasized here to point out the continuing importance and informational basis of library collections of objects—and certainly of special collections. The contrast with the electronic environment is striking. Because there will be no artifacts on the network—by definition, so to speak—there arises no physical evidentiary information to assist in the study of the texts themselves or to provide a history of their transmission. Information exists on the network as recorded bits of information collected as a digital object. It is easily copied. It also is easily modified. The great asset of electronic information is also its great liability.

Electronic Collections Are Not Special

There is, therefore, nothing "special" about an electronic collection. There are no artifacts to provide added value to the substantive information. There is no concept of rarity because any digital object may be quickly copied any number of times. There is no concept of origin, in the sense of one document being the more authoritative source for another. One copy is just as good as any other identical copy, or as the original. For example, the copy of a manuscript on floppy disk (assuming one does not change it) is as good as the hard-disk copy; there is no

priority. There is no need for protection against use: Use will not damage an electronic work. And it is important to recognize that there is no difference between important materials and less important materials in any of these respects. There are no special collections on the network.

Concepts of preservation and protection still exist.[6] They exist, however, in new terms for digital objects. In the short term, the medium the object resides on still needs to be protected and preserved (medium preservation). In the longer term, the digital object needs to be preserved against inevitable advances in technology, both hardware and software (technology preservation). And at all times in the electronic environment, the integrity or authenticity of the object needs to be guaranteed for the user to have assurance that the information is what it is expected to be (intellectual preservation).[7]

The computing term for the long-lasting quality of a digital object is *persistence*. In most cases, the decision of a library to provide access to an object—particularly, if it maintains a copy locally or provides full cataloging for it—is a decision that recognizes the desirability of persistence. In most cases, for electronic objects it will be a decision in favor of persistence because the decision to provide access will, in future, be necessarily linked to the decision to preserve the information.

Information "out there on the net" is not persistent until some institution assumes responsibility for its persistence, just as printed materials in the past have had no guarantee of survival until consciously acquired and preserved. The institution best prepared to take this responsibility remains the library. Sometimes it is loosely asserted that libraries will have a lesser role in the future because information will be available to everyone on the network. The fact remains that for information to be available for any meaningful length of time, someone has to select it and take responsibility for it, which

has been—and remains—the role of the library. Libraries continue in the paradigm learned in library school of acquiring (or linking to) information, organizing it, making it available, and preserving it.

Until now, the decision to provide access has been separate from the decision to preserve because, in fact, books can wait quite a while on the shelf before the preservation decision has to be faced. With digital objects, the decision must be faced from the moment of choice because digital information is volatile. From the beginning, therefore, electronic collection development decisions are decisions about the desirability (or otherwise) of persistence, or permanence.[8]

From the beginning, therefore, electronic collection development decisions are decisions about the desirability (or otherwise) of persistence, or permanence.

In the electronic environment, the decision to provide long-term access will become analogous to the decision to acquire, and it will mean making a commitment to long-term preservation. It should be noted that in this context there is no difference between what once might have been considered special collections objects and other more ephemeral items, whether temporary Web pages, e-journal articles, "texts" of authors writing in electronic form (perhaps even working texts from disks), or page images of decaying Dead Sea Scroll fragments that will become the primary record.

Special Collections on the Network

The role of special collections departments on the network is not yet fully clear, but some generalizations may be made. First and foremost, special collections departments will continue in their role of being the locus for the important artifact. Second, they will have an important role in placing digitized images of artifacts on the network.[9] Finally, they will have *no* distinctive electronic role within otherwise electronic libraries.

The special collections department will continue to be the part of the library that maintains and provides important artifacts of the human record. Existing collections can be expected to grow indefinitely simply by acquiring necessary and desirable materials from the preelectronic age. And despite some predictions, writing and printing of importance are going to be take place well beyond our lifetimes. Such activities will provide materials for traditional special collections oversight for some time to come. All these materials will need to be maintained and preserved for the indefinite future just as we always have supposed.

Placing digitized materials on the network already has become a requirement for present special collections departments. Rare book departments and treasure collections always have played a role in enhancing the prestige of their institutions. Many university administrators (and some librarians) already have seized on the public relations potential of placing digitized versions of attractive holdings on the network; it appears that some think that digitization represents all there is about electronic libraries. Too often, insufficient thought has been given to the infrastructure requirements of digitizing selected parts of collections, including costs, personnel, systems, and preservation; but this has not stopped the recent boomlet in digitizing projects characterized by scattered focus and varying standards. Still, they have had the advantage of helping train a number of library staff in the requirements of digitization and storage, including the niceties of resolution, compression ratios, image formats, and disk space.[10]

In the long run, when consistently organized and funded, digitizing projects will prove of real value for libraries and their users—and for many materials.

Access to their content will be improved while the materials themselves will be preserved. Rather as an exhibit lets patrons see an illuminated manuscript or Pound's marked-up manuscript of Eliot's *The Waste Land*, so digitized images can be placed online as a virtual exhibit; and they can be left online more or less indefinitely, an advantage over most artifactual exhibits. Accessibility of individual items can be improved for both pleasure and research, thus reducing the wear and tear on items that otherwise might be too popular to allow public handling. Thus, network surrogates can enhance the preservation of artifacts by allowing their study without actual physical handling. For some objects, the stress of digitizing will have to be balanced against the stress of handling, but for many objects the balance will be clear. In some instances, the digitizing actually may improve access, as in the case, for example, of the Beowulf manuscripts where digitizing using special lighting actually has made some parts of the text more readable than through handling the object itself.[11]

Unnecessary travel can be minimized if some part of scholars' needs can be satisfied by use of a well-digitized surrogate on the network. In other cases, the availability of a digitized collection can make a scholar's travel more efficient by allowing him or her to verify the presence of a document and to identify particular documents of interest before arrival at the actual site. Comparison of documents at geographically separated sites will be made much easier.

If digitizing is accompanied by other forms of editing, new forms of study will become possible. For example, formal markup (e.g., the Text Encoding Initiative [TEI], a form of the Standard Generalized Markup Language [SGML])[12] will allow computer manipulation of texts and testing of hypotheses in ways not possible in a precomputer environment.

All these digitized uses raise again the question, not to be answered here, of who the constituency for the rare book library is—members of the local institution or the scholarly community at large. For libraries with important special collections departments, there always have been the tensions between local needs and local funding on the one hand, and national prestige and extramural users on the other. The network will only heighten these tensions.

Special collections departments no longer will collect distinct kinds of information within the digital research library. As described above, there is no distinctive activity for a special collections department in the selection, organization, use, or preservation of digital information. The same activities will take place for the identification and long-term maintenance of a digital textbook as would take place for preservation of a human genome database or of a hypertext novel by Robert Coover.

Thus, network surrogates can enhance the preservation of artifacts by allowing their study without actual physical handling.

Thus, one implication is that, to some extent, existing special collections departments may become museums of the book, not of course in the debased sense of mausolea but, rather, as places where objects are studied rather than texts. Such a role would be mitigated partially by digitization activities to make textual collections more widely available. Special collections, in fact, would then become a bridge backward in time—a "legacy system," in computer jargon—as they continue to fulfill their role as a repository of original documents of the human record.

The Research Library As a Special Collection

Alternatively, of course—and a preferable way of seeing our future—it could be said that the entire research library will become

a special collection, preserving the electronic human record for use into the indefinite future. (Rare book librarians might call this approach: "If you can't beat 'em, join 'em.") In this view, there are some interesting implications for special collections staffing.

The historical curatorial skills will remain important in research libraries.

The considerations involved in selecting important originary digital works will be the only distinction between what now are called special collections curators and research library bibliographers. After the object is selected, whether it is the electronic working papers of a Nobel prize winner or the next electronic novel by Danielle Steele, the library activities of cataloging, authentication, preservation, and presentation will be the same.

The distinctive characteristic for the digital curator of special collections may be a special understanding of intellectual property and of how copyright will color the passing of ownership or license from a private party to the library. What now is thought of as normal or trade materials will have their electronic analogy in the routine acquisition of digital objects through standard purchase, lease, or license agreements. The digital curator, on the other hand, will sniff out particular digital objects of importance and individually negotiate rights to them with their owner. Electronic copyright law may become an important hiring qualification in special collections.

So far, the question has been begged of what is an "important" digital work. In the artifactual environment, the concept often is clear, though sometimes controversial. The criteria often change. The hard-copy concept of importance is affected by changing views of culture, for example, as in the rise of gender studies or of popular culture

and ephemera. Still, importance is given a concrete foundation by artifactual matters such as edition, printing state, binding, age, and demonstrable provenance. Thus, in a broad sense, we know an important artifactual work when we see one, which is what allows the antiquarian book trade to thrive.

But what is an important digital work? The hypothesis so far has been that everything about research library activities involving digital works is the same except for the selection process. And the selection process is different only for "important" works. Perhaps what will be thought of as "important" digital works will be those desirable digital objects for which the copyright issues are troublesome. They will not be any rarer than anything else, they will not be any more fragile, they will not be any more esthetic, and they will not provide any material supporting evidence for the substance of their information. It may be the concept of scarcity—that is, the difficulty of acquiring rights—that provides the sense of importance or of value.

Saying that a digital work is important because it invokes the specialized skill of a digital copyright curator is not yet a very crisp or satisfactory working definition. The uncertainty of this definition also may be heuristically helpful in understanding why special collections *as such* will not really exist on the Internet.

However, we may be confident that there will continue to be a role for special collections in the networked environment. The historical curatorial skills will remain important in research libraries. These are the skills for understanding what is important in the literature, for knowing who the important players are, for asserting what a given library should take responsibility for, and for negotiating with private parties.

In his 1990 Malkin lecture, Terry Belanger pessimistically described the role of special collections as leading the march to the dumpster as research library

shelves get fuller.[13] Instead, the role of special collections should be to lead the march to the network. In rejoining their main library colleagues, special collections librarians will be leading themselves back into the mainstream of the research library. This will be of enormous benefit to their colleagues, to themselves, and to library patrons now and in the future.

Notes

1. This article refers to "books" as shorthand for the many different kinds of artifactual objects in libraries and special collections, including manuscripts, broadsides, pamphlets, journals, sheet music, recordings, photographs, and the like. The general point*, mutatis mutandis,* holds for all of them.

2. See the dust jackets in any contemporary bookstore. For the scene 200 years ago, see George Crabbe (1754–1832), "The Library" (1808):

> ...Lo! all in silence, all in order stand,
> And mighty folios first, a lordly band;
> Then quartos their well-ordered ranks maintain,
> And light octaves fill a spacious plain.
> See yonder, ranged in more frequented rows,
> A humbler band of duodecimos.

3. For example, see D. F. McKenzie on the shift in Congreve editions from quarto to octavo. The new format and the addition of stage directions, italic, and decorations signaled a change in how author and publisher intended the work to be taken by its audience, and thus the shift in sensibility now referred to as the shift from Restoration to Georgian; "Typography and Meaning: The Case of William Congreve," in *Buch und Buchhandel in Europa im achtzehnten Jahrhundert*, eds. Giles Barber and Bernhard Fabian (Hamburg, 1981), 81–125. McKenzie's work is cited in Robert Darnton, *The Kiss of Lamourette* (New York: W. W. Norton, 1990), 132 and 184.

4. For examples, see Roger Chartier, *The Order of Books* (Stanford, Calif.: University Pr., 1994), and Anthony Grafton, "Is the History of Reading a Marginal Enterprise? Guillaume Budé and His Books," *Papers of the Bibliographical Society of America* 91 (June 1997): 39–57. See especially Roger Stoddard, ed., *Marks in Books, Illustrated and Explained* (Cambridge: Houghton Library, 1984), and the December, 1997, *Papers of the Bibliographical Society of America* devoted to the 1997 Marks in Books Conference of the BSA. For a more general introduction, see Robert Darnton, "What Is the History of the Book?" in *The Kiss of Lamourette* (New York: W. W. Norton, 1990).

5. The classic example is Charlton Hinman's *The Printing and Proof-reading of the First Folio of Shakespeare,* 2 vols. (Oxford: Clarendon Pr., 1963). A briefer version is his introduction to the Norton Facsimile, *The First Folio of Shakespeare* (New York: Norton, 1968). For more on analytic and descriptive bibliography in this tradition, see the works of G. Thomas Tanselle, for example, in almost any of the annual volumes of the *Papers* of the Bibliographical Society of the University of Virginia.

6. *Preserving Digital Information: Final Report and Recommendations* (Washington, D.C.: Commission on Preservation and Access/Research Libraries Group, 1996). Available in print and on the Web at: <URL:http://www.rlg.org/ArchTF/>.

7. "In this new world, preservation means copying, not physical preservation." Michael Lesk, *Preservation of New Technology: A Report of the Technology Assessment Advisory Committee to the Commission on Preservation and Access* (Washington, D.C.: Commission on Preservation and Access, 1992), 13; also at: <URL:http://clir.stanford.edu/cpa/reports/lesk2.html>. See also Clifford A. Lynch, "The Integrity of Digital Information: Mechanics and Definitional Issues," *Journal of the American Society for Information Science* 45 (1994): 737–44; Peter S. Graham, *Intellectual Preservation: Electronic Preservation of the Third Kind* (Washington, D.C.: Commission on Preservation and Access, 1994); also at: <URL:http://aultnis.rutgers.edu/texts/cpaintpres.html>.

8. Patricia M. Battin noted the need for preservation decisions at digital creation in the 1992-93 *Annual Report* of the Commission on Preservation and Access, p. 3. The archiving community has emphasized "life-cycle" treatment of records to assure preservation, and this approach is

now appearing in digital archiving discussions; see e.g. Neil Beagrie and Daniel Greenstein's upcoming *Digital Collections: Developing a Strategic Policy Framework for the Creation and Preservation of Digital Resources* (London: Arts and Humanities Date Service, draft of March 9, 1998).

9. Abby Smith has summarized the value of the acces and preservation role in "Special Collections Stake Their Claim in the Electronic Age," *CLIR Issues,* (March/April 1998): 1.

10. For an outstanding introduction to these topics, see Anne R. Kenney and Stephen Chapman, *Digital Imaging for Libraries and Archives* (Ithaca, N.Y.: Cornell Univ. Library, 1996). Proposal responses by many libraries to the Ameritech/LC grant sponsorship are examples of what is noted in this paragraph. In contrast, see the recent statement by the Society of American Archivists, "The mere potential for increased access to a digitized collection does not add value to an underutilized collection. It is a rare collection of digital files indeed that can justify the cost of a comprehensive migration strategy without factoring in the larger intellectual context of related digital files stored elsewhere and their combined uses for research and scholarship," *Statement on the Preservation of Digitized Reproductions* (June 9, 1997), <URL:http://www.archivists.org/governance/resolutions/digitize.html>.

11. Kevin Kiernan, "Digital Preservation, Restoration, and Dissemination of Medieval Manuscripts," (1993), <URL:http://www.uky.edu/~dhart/dhart.html>, and "The Electronic Beowulf" (The British Library, 1995), <URL:http://www.uky.edu/~kiernan/BL/kportico.html>.

12. See the pilot project developed by Wendell Piez and others at the Center for Electronic Texts in the Humanities (CETH) of Rutgers and Princeton, "The William Elliot Griffis Collection Electronic Access Project," (1996), at <URL:http://www.ceth.rutgers.edu/projects/griffis/project.htm>. The project uses SGML DTD's both for Encoded Archival Description (EAD) and the Text Encoding Initiative (TEI).

13. Terry Belanger, *The Future of Rare Book Libraries* (Columbia University School of Library Service, Book Arts Press: 1991 Malkin Lecture; text available from Dec. 16, 1991, archive of *ExLibris*, message from: terry@cunixa.cc.columbia.edu, subject: Malkin Lecture).