Selection for Preservation in the Digital Age

An Overview

Janet Gertz

In this article, I address three sets of issues. First, is digital conversion a preservation technique or is selection for digitization fundamentally an issue of access? Second, how does the process of selection for digitization differ from selection for traditional preservation activities? What selection criteria apply? Finally, what effect might digitization have on preservation as a field?

In this paper, I consider three issues: whether conversion to digital form is a preservation action, the contrasts between selection for digital conversion and selection for traditional preservation, and the potential effects on the field of preservation. All of these issues are under active debate by, among others, Atkinson (1998), Conway (1996b), Hazen et al. (1998), and Smith (1999). While consensus is growing, many points remain unsettled.

Is Digitization a Real Option for Preservation?

It is a given that in order to achieve preservation, we must provide a long-term version of at least the intellectual content of the item. Ideally we preserve the original object itself, appropriately repaired and properly housed. Driven by necessity in the form of irreversible deterioration, we produce surrogates. While it is of course impossible to guarantee permanent survival (O'Toole 1989), preservation relies on the use of stable media with long life expectancy, properly made and properly stored, to prolong the existence of the information.

Once created, permanent and durable paper copies and preservation-quality microfilm will endure for hundreds of years, barring disaster or vandalism. To date no one can prove that any digital version will survive and still be accessible beyond a few decades, despite much talk about migration and emulation, especially considering the repeated intervention these will require. Further, the accuracy and authenticity of a digital version may be open to question. Was the original object accurately and completely represented in the digital version at the time of image capture? Can we be assured of its authenticity over time—that is, can we be sure that it is still complete and has not suffered undocumented change? Lacking agreed-upon mechanisms for this assurance, and lacking longevity, digital copies alone cannot constitute preservation.

It is also a given that preservation without access is futile. Digitization offers enhanced, wider, easier access than microforms or photocopies, and it can capture color, sound, movement, and other features that traditional preservation reformatting has not handled well. The term "hybrid approach" (Willis 1992) expresses the idea that digitization in combination with traditional preservation activities provides a way to accomplish both sides of the preservation/access dyad: longevity via traditional means and

Janet Gertz (gertz@columbia.edu) is Director for Preservation, Columbia University, New York, N.Y.

This article is a revised version of the paper of the same title presented June 26, 1999, in New Orleans at the ALCTS-sponsored PARS program "Selection for Preservation in the Digital Age."

Manuscript received August 30, 1999; accepted for publication November 12, 1999.

improved use through digital means. We now have the option to microfilm an item, then scan the film (Conway 1996a); scan the item then produce computer output microfilm or print out to acid-free paper instead of making photocopies or copyflo (Kenney 1997); or we can scan the original item, retain it, and use the online version as a facsimile or surrogate to protect the original from unnecessary handling. All of this forms part of a continuum with preservation at one end and digitization purely for access purposes at the other. The issue for selection is deciding when best to employ digitization.

Is the Process of Selection for Digitization Significantly Different From Selection for Traditional Preservation Activities?

The process of selection for traditional preservation reformatting and the criteria on which decisions are based are well established (see Williams and Lunde 1997 for an historical overview). The decision process is not linear. No single selection criterion suffices; it is only valid in combination with others, as the decision-maker steps through a series of complex, interconnected questions where each answer influences the others. Further, answers to these questions are situational. Given my institution's specific mission and history, is this item old enough, useful enough, important enough to keep? Given the size of my preservation budget and the number of items that need care, is it damaged or endangered enough to warrant expenditure? Selectors at different institutions frequently give different answers when faced with copies of the same book.

Physical condition drives traditional preservation decision-making. An item added to the collection sometime in the past is now in fragile, damaged, or threatened physical condition. We evaluate its value for continuing scholarship. Decisions follow to determine how preservation should and can be achieved, based on the item's value, its physical properties, and the nature of its current and anticipated use, as in the following criteria:

- Is the item or collection damaged or endangered?
- Does it have sufficient enduring value to justify preservation? This can include but is not limited to: artifactual features, e.g. bindings, illustrations; uniqueness or historic importance; distinguished broad or deep long-term intellectual content, with potential long-term value for teaching or research; consonance with the mission of the institution; and contribution to or support for historically important areas at the institution.
- Which preservation options are available, given the physical nature of the item or collection, and its current and predicted future use? Can we repair it? If not, can a preservation copy successfully capture its content, and support current and predicted future use? Are there factors like use of color, poor contrast, or missing pages that might make traditional reformatting inappropriate?
- What is the cost of each preservation option? Which best matches the monetary and intellectual value of the item or collection?

Many of the same criteria hold for selection for digital conversion but with changed emphasis. The fundamental difference is that the selection process most often starts from a desire for better access rather than due to physical deterioration. This is only to be expected with a technology whose primary characteristic is ease and breadth of access rather than longevity. The primary criterion of enduring value is the same for preservation and digital conversion, but user demand gains in importance, and issues such as intellectual property rights and technological potential receive more attention than is the case with traditional methods.

As with traditional preservation selection, identifying an item or collection as a possible candidate for digital conversion means determining both whether the technology can do what is wanted and whether it is worth doing. And like selection for traditional preservation, selection for digital conversion also functions through local interpretation of general principles. There are no absolute answers, only questions that must be answered within the local context. A number of libraries and organizations have developed criteria for selection for digital conversion, among them Arizona State University Library (1998), Columbia University Libraries (1997), Library of Congress (1999), Smithsonian Institution Libraries (1997), Society of American Archivists (1997), University of California (1997), and University of Illinois (1998). Fuller discussion and quotations from a number of them are available in Gertz (1998); see also Hazen et al. (1998). All of the authors of these documents agree fairly closely on the criteria for selection that they propose, and they share the goal of selection to match the strengths of digitization with the materials it handles best.

The most often cited criteria can be summed up as follows:

- Does the item or collection have sufficient value to and demand from a current audience to justify digitization?
- Do we have the legal right to create a digital version?
- Do we have the legal right to disseminate it?

- Can the materials be digitized successfully?
- Do we have the infrastructure to carry out a digital project?
- Does or can digitization add something beyond simply creating a copy?
- Is the cost appropriate?

Value and Demand

The first criterion is the same basic question that holds for selection of new acquisitions, collection review, or preservation—whether the content of the materials has enduring value. In the case of digitization, a high level of current demand for the materials or likeliness to undergo significant use once they are digitized adds weight to the decision. Among the selection guidelines mentioned above, phrases such as these appear frequently:

- Does the proposed item or collection have active current users?
- Is there greater demand than can be served by the original or a traditional type of copy?
- Does it support high priority activities such as teaching of core courses that have large enrollments?
- Is it marketable to a group of specialists widely dispersed who all need access?
- Do limitations on handling of fragile or valuable originals create a source of demand for high quality surrogates?
- How does it fit with other materials on the same subject?
- Does it help build a distributed online collection?

Materials that are truly in high demand tend to be converted by commercial agencies, because they offer a likely source of profit, while libraries and other cultural institutions tend to focus more on the unique items that they hold and on the lesser-used materials that do not attract commercial interest. Brittle books (many unused for the past several decades) raise particular questions. A great deal of preservation energy goes into microfilming very low-demand brittle books that are selected specifically because they have potential research value for some future scholar—and so ought to be preserved—but are not high priority for many (if any) current researchers. They can reasonably be put on a successful storage medium like film, even though it is a slow, awkward access medium.

Projects like Making of America (Crist and Price-Wilkin 1996) move brittle books into the digital arena, raising the question of whether or to what extent digitization is appropriate for such materials. Atkinson (1996) articulates the concept of the distance between the scholar and information. The book on the shelf is a certain physical distance from the scholar; microfilm moves it further away physically by interposing the need for equipment and psychologically through dislike of the medium. Digitizing and mounting the images on the Web brings the content much closer. When is it appropriate to digitize brittle books as well as—or, like the University of Michigan, instead of—microfilming when the immediate demand for such materials is low and preservation is the primary motive for reformatting?

Intellectual Property Rights

Once selectors establish that the materials merit digitization, they must then determine whether they have the legal right to make digital copies and disseminate them, and if not, whether they can get clearance from the rights holder. Preservation has largely been focused on microfilming printed matter that is no longer under copyright, and in any case institutions have filmed to replace or protect damaged items for local use or to provide copies clearly for research purposes—all legitimate under copyright law. What is legitimate for digital versions is less obvious. The 1998 changes to the copyright law permit creation of digital preservation copies, but whether and how access to them may be provided remains open to interpretation (ALA Washington Office 1998).

Digitization increasingly focuses on unique archival and visual materials, many of them unpublished and often very desirable for nonresearch applications. We run into unclear histories of ownership and multiple layers of authorship. Intellectual property issues become murkier as online versions join the mix because they are universally accessible and can be infinitely copied and altered without knowledge or permission of the rights holders. To protect intellectual property from unauthorized use, some libraries allow access for institutional affiliates only. Other institutions mount low-resolution images that are inappropriate for printing but which may also be too low for serious research use. The tension between protecting intellectual property rights and broadening access has become a serious issue for the selection process. Adding fuel to the fire, it seems possible that converting these materials into digital form might actually be a profitable activity—which is of course when rights holders start paying attention.

Technical Feasibility

As with traditional reformatting, we must ask whether the technology can in fact do what we need. Can the materials be captured adequately? How close can the digital version come toward representing the full content of the original? Can digital versions be made that will serve most if not all of the same functions as the original? If not, can the digital version fill enough functions to justify its creation, and how will the others be

provided for in future? Further, can adequate digital versions be created without damaging the originals? An informed selection decision can only be made with a solid understanding of whether quality images can be captured, stored, delivered legibly and conveniently through commonly used equipment and software.

A great deal of work is going into defining quality standards. Ideally, we aim for high quality in most of the digitization we carry out, especially for preservation purposes. However, it is also true that, as with microfilming, we will not always be able to meet guidelines for highest quality due to limiting circumstances. For instance, we may be considering a fragile item that must be retained in original form and cannot be laid flat or disbound. Or, poor quality older film may be the only version of the item left to scan if the original object is badly deteriorated or no longer exists. Is a digital version the best option under the circumstances? When is a lower-quality digital copy better than no copy at all?

Infrastructure

Rational selection decisions cannot be made without understanding whether a project is feasible in terms of the institution's digital infrastructure. In the past, this has meant finding a good filming bureau and learning to catalog microfilm. Digital conversion is inextricably and essentially bound up with levels of work never needed for traditional reformatting: determining the appearance of the page on screen; navigational tools; user interfaces; structuring and indexing; and generating metadata to record information about images and make them usable. Now we must determine not only whether we can adequately prepare the materials and capture their content digitally, but also whether we can create metadata and manage files, handle intellectual control, and make the digital images functional. All of these tasks are basic to an effective product, and the resources needed to achieve them must be factored into selection.

Added Value

The next issue is whether digital conversion can do more than traditional reformatting—whether it can add value rather than simply creating a copy. Some value-added options relate to image quality: can and should images be manipulated to make them more legible than the original items, for instance, by removing discoloration and stains? This has direct bearing on the perceived accuracy and authenticity of the digital version compared to the original, but it may make the digital images substantially more useful. When might it be appropriate and when not?

Other options involve enhanced description and searching, for instance, item-level cataloging for visual materials, or optical character recognition (OCR) to make digital images of texts fully searchable. Simple bit-mapped images of text pages are only marginally more accessible than microfilm, and in fact Smock (1995, B2) has asked, "Why bother to convert a text to electronic form unless one can do more with it than with the printed version?"

Costs

Once we have ascertained that digitization is appropriate and feasible, we return to the familiar issue of money, in this case evaluating the cost of digitization compared with other preservation options. It has been the experience at my institution that digitization is more expensive than microfilming when all costs for image capture, preparing the materials, file management, Web interface design and programming, and intellectual control activities are counted (see appendix).

While hard numbers are not easily available and are in any case difficult to compare across institutions and across projects, Conway (1996c) found that when scanning preservation microfilm, 32% of the costs went to indexing, i.e., the tasks needed to tie the file to the correct page number and place in the structure of the book. Neither Conway, Kenney (1997), nor Chapman et al. (1999) include in their costs any higher-level metadata or cataloging work. If one adds in OCR and further "value added" work involved in making digital images searchable and capable of manipulation, it all adds up to much more than the cost of traditional bibliographic control for microfilm, which we have long known constitutes up to 30% of project costs by some calculations (McClung 1986; Kantor 1986). How much intellectual value and potential for future use does a volume or a collection need to have, or how much more searching and manipulation need to be added online, to justify the expense?

There is also of course the question of whether funding is available. On rare occasions an institution may recoup its expenses by creating a marketable product or by contracting with a publisher who will pay royalties for digital images of high-profile and rare items. But most libraries will be converting materials with little money-making potential. Most project funding comes from grant agencies or cooperative programs, and part of the whole selection process lies in deciding when participation in externally funded projects matches the institution's overall needs.

Developing Strategies and Priorities for Digitization

All of this said, it is clear that each institution must develop strategic plans and priorities for digital conversion grounded in its mission and goals or risk enormous waste of human and financial resources on projects that use digitization when other technologies

would be more appropriate, or that are poorly designed, or that focus on the wrong materials.

A brief review of selection strategies from two institutions demonstrates how institutional missions lead to divergent priorities for digitization. First, consider the digitization guidelines of the National Agricultural Library (NAL 1995). The NAL is a national library with responsibility as the repository of record. As such, librarians there have developed guidelines that focus on documenting the history of agriculture and the history of the government agencies they serve, through digitization of both brittle printed works and unique items. They give preference to works of historical value that contribute to an understanding of the history of agriculture and fill gaps in the history of the Department of Agriculture, and they digitize materials that are in the core area of their collecting responsibility. High demand is not the primary factor here; rather, it is NAL's role as the national archive for a specific subject area and the need to preserve and make widely accessible the central materials in that area.

An alternative route is taken by the Beinecke Rare Book and Manuscript Library at Yale, in a recently published description of its institutional strategy for digitization projects appropriate for a special collection holding primary materials and rare books. Bouche (1999, 4) outlines this strategy: "The library sought to clarify a sustainable strategy for an ongoing commitment to a digital component within its established operations. . . . Given . . . that the library could neither afford nor programmatically justify any attempt to scan everything that might be of potential interest to scholars . . . , where ought the library to place its emphasis to yield the maximum benefit?" Librarians there have focused on two types of projects they believe make the best use of the unique strengths of digital technology for their purposes. First are projects to add value to primary research materials "where a case can be made that the study and interpretation of the documents could materially benefit from conversion to digital form, allowing flexibility in viewing and assessing . . . in a manner that cannot be derived readily from a good quality microfilm, photocopy, or photographic duplicate, or possibly even from close inspection of the original itself." The second focus is visual materials, because of "the overall benefits in access, preservation of fragile originals, and overall reader services" (Bouche 1999, 14–15).

What Effect Is Digitization Likely to Have on Preservation?

Digital conversion allows us to satisfy our clients better while carrying out preservation because we can actually create a use medium they like and want to use. This is of course the whole reason we are concerned with digitization in the first place. It has the potential to please our patrons, offers them capacities for research not available before, and broadens the range of media we can deal with because it can capture text, color and continuous tone images, sound, and movement. Instead of "just" trying to solve the brittle paper problem, we now have the potential to convert other media we have avoided for many years, and to do it with a technology that users actively like.

Digital conversion will make us review and refine traditional procedures to improve them and to make the interaction of digital and traditional technologies easier and better. Chapman et al. (1999), Kenney (1997), Conway (1996a), and others suggest modifications to microfilming procedures in order to assure that microfilming and digitization will interact better in future, with a better quality product in both cases. Some of those suggestions involve compromises with current microfilming specifications. If compromise cannot be avoided, then the preservation community must decide when it is appropriate. As we work with audio and visual materials, we will be constantly moving between what we traditionally have been able to achieve and what digitization offers, and making sure we continue to value the core concept of preservation—the primacy of quality and longevity—in both traditional and new technologies.

The current realities of digital conversion will also make us rethink the issue of achieving high image quality by employing capture processes that harm or destroy the original object in order to preserve its content. The clearest scanned images of book pages are currently achieved most cheaply by flatbed scanning from cut pages (Chapman et al. 1999). We need to weigh obtaining a better-quality product for less money against destroying the original object. This is not a new dilemma; see for instance Ogden (1989). Microfilming has always caused some wear and tear on materials, and there was a time when many of us routinely disbound volumes and discarded them after filming. Readers objected strongly not merely because they dislike microfilm, but through genuine concern about the destruction of the original volumes, as stated in the Modern Language Association statement (1995) and perhaps most strenuously expressed by Tanselle (1989, 44), who speaks of "the unnecessary destruction of books in the name of textual 'preservation.'" Although we have developed better ways to limit filming damage as far as possible, some libraries still face hostile inquiries about withdrawal of materials after filming (Singer 1998). Will readers react less angrily if digitization requires destruction, given that it is an access medium many of them relish? Or do we decide to live with lower

image quality or more expensive digital capture in order to keep originals intact? We can hope that improvements in technology will eliminate the quandary soon.

Digitization has the potential to bring in more resources, because it is a high status activity, and administrators, grant agencies, and donors are perhaps more willing to pay for it than for preservation. On the other hand, higher costs mean that resources will not stretch as far as we would like. We must factor in the costs of indexing, metadata creation, development of navigational aids, Web site design, and the whole package that makes the digital copies useful and accessible to the viewer. Unlike microfilming, where bibliographic control can reasonably be segmented off from image capture, digital image capture is intimately bound up with creation of operational, structural, and administrative metadata. Essential information must be recorded about the conditions and mechanisms of image capture and about the innate structure of the object (page sequences, chapter divisions, grouping of archival materials, and so forth) to permit navigation among related files. Preservation experts must become as knowledgeable about metadata creation as about creation of the images themselves (RLG Working Group 1998).

Digitization absorbs a great deal of managerial attention, which translates into money. It complicates decision-making because there are so many variables: whether to create—and whether it is possible to create—full online facsimiles or else some level of lower resolution index or reference image; decisions about image quality, tonality, enhancement of images; whether to create searchable text and how, what level of complexity of metadata to create, what sort of search tools and user interface, Web site design, what storage media to use—the list goes on and on. It also takes time and energy to develop the necessary infrastructure of guidelines and procedures and tools like Requests for Proposals and contracts, to learn to scan to high standards, and to identify and cultivate vendors.

Digital conversion can divert attention and resources away from preservation to purely access projects. Providing digital resources to library users has become an essential service, both via conversion and through purchase or subscription. Administrators are desperately looking for lower priority operations whose funds and staff can be moved to support digital resources. The strengths of digitization come to the fore in displaying illustrated and visual materials, in providing ready access to high demand materials, and in creating new research tools by combining scattered resources. Digital conversion projects of this type usually have little to do with preservation; rather, they are almost purely access-driven. Do we redefine preservation to justify our roles in such projects or spread our time over a broader area of responsibility beyond preservation?

Finally, digital conversion will increase the amount of preservation that is needed because we will also need to preserve the digital resources we are producing. Whether we carry out digital conversion for preservation or other purposes, there is always an eventual preservation angle. Using digitization to create access to materials of long-term value calls for genuine commitment to preserving the digital files, to development of the infrastructure necessary to preserve those files routinely over the long-term. It is in preservation's best interests to be involved at the selection and digital conversion step to assure that images are of appropriate quality, stored on appropriate media, and accompanied by appropriate metadata.

Conclusion

Digitization has the potential to do so much, and there is a great temptation to rush off to create the digital resources we know are possible. Careful decisions and a great deal of money will be needed if we are to do it well. The wonderful capacity for searchable text and identification of images across the Internet can only be bought at a very high price. Digitization gives us a way to make preserved items infinitely more immediate and available to anyone connected to the Internet, not just scholars who can visit our library or borrow a copy of our microfilm. It gives us options far beyond what we had before, and it allows us to make use of a medium that is not merely user-friendly but actively sought by users. In many senses, we cannot afford not to digitize. The essential thing is to use digitization to its, and our, greatest advantage, and to use traditional techniques and digital conversion together to better effect the preservation of our materials by choosing carefully what is most appropriate for the materials and to maximize the strengths of both analog and digital technology.

Works Cited

ALA Washington Office. 1998. Library preservation: Changes incorporated in H.R. 2281. The digital millenium copyright act of 1998 (PL 105-304). Prepared by Carol Henderson. Accessed Nov. 12, 1999, www.ala.org/washoff/preservation.html.

Arizona State University Library. 1998. Digitization program pilot project application. Unpublished.

Atkinson, Ross. 1996. Rationality and realpolitik: Prospects for cooperative collection development in an increasingly networked environment. In *Scholarship in the new information environment: Proceedings from an RLG symposium held May 1–3, 1995, at Harvard University,* ed. Carol Hughes, 29–31. Mountain View, Calif.: Research Libraries Group.

Atkinson, Ross. 1998. Managing traditional materials in an online environment: Some definitions and distinctions for a future collection management.

Library Resources & Technical Services 42:7–20.

Bouche, Nicole. 1999. *Digitization for scholarly use: The Boswell Papers Project at The Beinecke Rare Book and Manuscript Library.* Washington, D.C.: Council on Library and Information Resources. Also available at: www.clir.org/pubs/reports/pub81-bouche/pub81text.html.

Chapman, Stephen, Paul Conway, and Anne Kenney. 1999. *Digital imaging and preservation microfilm: The future of the hybrid approach for the preservation of brittle books.* Washington, D.C.: Council on Library and Information Resources. Draft version, accessed Nov. 12, 1999, www.clir.org/cpa/archives/hybridintro.html.

Columbia University Libraries. 1997. Selection criteria for digital imaging projects. Accessed Nov. 12, 1999, www.columbia.edu/cu/libraries/digital/criteria.htm.

Conway, Paul. 1996a. *Conversion of microfilm to digital imagery: A demonstration project.* New Haven, Conn.: Yale Univ. Library.

———. 1996b. *Preservation in the digital world.* Washington, D.C.: Council on Library and Information Resources.

———. 1996c. Yale University Library's Project Open Book: Preliminary findings. D-Lib Magazine 2. Accessed Nov. 12, 1999, www.dlib.org/february96/yale/02conway.html.

Crist, Margo, and John Price-Wilkin. 1996. The access context. In *Selecting library and archive collections for digital reformatting: Proceedings from an RLG symposium held November 5-6, 1995, in Washington, D.C.*, p. 29–38. Mountain View, Calif.: RLG.

Gertz, Janet. 1998. Selection guidelines for digitization. *Guidelines for digital imaging: Papers given at the Joint National Preservation Office and Research Libraries Group and Conference in Warwick, 28th-30th September 1998*, 11–19. London: NPO. Also available at: http://lyra.rlg.org/preserv/joint.

Harris, Carolyn, Carol Mandel, and Robert Wolven. 1991. A cost model for preservation: The Columbia University Libraries' approach. *Library Resources & Technical Services* 35: 33–54.

Hazen, Dan, Jeffrey Horrell, and Jan Merrill-Oldham. 1998. *Selecting research collections for digitization.* Washington, D.C.: Council on Library and Information Resources. Also available at: www.clir.org/pubs/reports/hazen/pub74.html.

Kantor, Paul. 1986. *Costs of preservation microfilming at research libraries.* Washington, D.C.: Council on Library Resources.

Kenney, Anne. 1997. *Digital to microfilm conversion: A demonstration project, 1994–1996.* Ithaca, N.Y.: Cornell Univ. Library. Also available at: www.library.cornell.edu/preservation/pub.htm.

Library of Congress. Preservation Reformatting Office. 1999. Selection criteria for preservation digital reformatting. Accessed Nov. 12, 1999, http://lcweb.loc.gov/preserv/prd/presdig/presselection.html.

McClung, Patricia. 1986. Costs associated with preservation microfilming. *Library Resources & Technical Services* 30: 363–74.

Modern Language Association. 1995. Statement on the significance of primary records. *Profession* 95: 27–28.

National Agricultural Library, Selection Task Force of the Electronic Preservation Committee. 1995. Selection criteria and guidelines: Selecting materials for digital preservation at the National Agricultural Library. Unpublished.

Ogden, Barclay. 1989. *On the preservation of books and documents in original form.* Washington, D.C.: Commission on Preservation and Access.

O'Toole, James. 1989. On the idea of permanence. *American Archivist* 52: 10–25.

RLG Working Group on Preservation Issues of Metadata. 1998. Final report. Accessed Nov. 12, 1999, www.rlg.org/preserv/premeta.html.

Singer, Mark. 1998. Talk of the town: Missed opportunities dept. Did the New York Public Library let some history slip through its fingers? *The New Yorker* Jan. 12: 29.

Smith, Abby. 1999. *Why digitize?* Washington, D.C.: Council on Library and Information Resources. Also available at: www.clir.org/pubs/reports/pub80-smith/pub80.html.

Smithsonian Institution Libraries. 1997. Guidelines for developing an SIL digital project. Unpublished.

Smock, Raymond. 1995. What promise does the Internet hold for scholars? *Chronicle of Higher Education* 42: B2.

Society of American Archivists. 1997. Statement on the preservation of digitized reproductions, approved by the SAA Council June 9, 1997. Accessed Nov. 12, 1999, www.archivists.org/governance/resolutions/digitize.html

Tanselle, G. Thomas. 1989. Reproductions and scholarship. *Studies in Bibliography* 42: 25–54.

University of California, Preservation Program. 1997. University of California selection criteria for digitization. Accessed Nov. 12, 1999, www.library.uscb.edu/ucppp/digselec.html.

University of Illinois Library, Digital Imaging and Media Technology Initiative. 1998. Questions to consider before beginning an image database project. Accessed Nov. 12, 1999, http://images.library.uiuc.edu/20questns.html.

Williams, Sara, and Diane Lunde. 1997. Preservation and collection development in academic libraries of the United States: Joint history and future prospects: A review article. *Advances in Librarianship* 21: 73–89.

Willis, Don. 1992. *A hybrid systems approach to preservation of printed materials.* Washington, D.C.: Commission on Preservation and Access. Also available at: www.clir.org/pubs/reports/willis/index.html.

Appendix
Cost Elements of Digital Conversion

We have not yet formally established costs for digital conversion efforts at Columbia, while microfilming costs have been tracked in detail for more than ten years (see Harris et al. 1991). In computing current costs for microfilming projects, we include the activities from initial identification of materials to be reformatted through delivery of the finished product to the user:

- condition survey and inventory of materials
- retrieval from the stacks and circulation tracking of materials
- professional bibliographer title-by-title review for retention and preservation decisions
- bibliographic searching for existing preservation-quality microfilm and for cataloging copy
- page-by-page collation and interlibrary loan or other means of acquiring missing parts
- target creation
- shipping to and from the filming vendor where image capture and creation of copy negative and positive film are carried out
- upgrading or creation of bibliographic records
- technical and frame-by-frame quality control inspection of the film, and arranging for any corrections
- distribution of the positive to the microform reading room and the negatives to off-site storage
- disposition (conservation, boxing, reshelving, or withdrawal) of the original materials.

The process of digital conversion includes all of the above tasks except bibliographic searching for existing microfilm and creation and distribution of film copies. To the above tasks are added:

- initial sampling and benchmarking to determine appropriate capture specifications and decisions on resolution, size, and quantity of derivatives
- creation in-house or by the vendor of derivatives
- creation partly by the vendor and partly in-house of operational, structural, and administrative metadata
- programming and creation of html pages for display and navigation
- quality control not only of each image but also of links to assure that all files open to the correct image
- general file management, storage, and backup.

Where possible, activities are automated, but the additional steps and the involvement of programmers and other technical staff inevitably add significant expense above the costs for tasks shared with microfilming projects.