

# The Changing Role of Special Collections in Scholarly Communications

Donald J. Waters, Senior Program Officer, Scholarly Communications,  
Andrew W. Mellon Foundation

*Presented at the ARL-CNI Fall Forum on "An Age of Discovery: Distinctive Collections in the Digital Age," Washington DC, October 14, 2009*

In October 1995, I was co-chair of the Task Force on the Archiving of Digital Information. I stood before the assembled membership of ARL and reported dutifully on the progress of the task force.<sup>1</sup> I noted in my talk then that William Safire had recently devoted his wonderful "On Language" column in the *New York Times Magazine* to the topic of kids' slang. He advised that "if you want to stay on the generational offensive, when your offspring use the clichéd 'gimme a break,' you can top that expression of sympathetic disbelief with 'jump back' and the ever-popular riposte 'whatever.'" However, he also noted that some expressions, such as "I'm outta here" or "I'm history," had become very much dated. Quoting from a study of slang, Safire pointed out that "I'm history," is "a parting phrase modeled on an underworld expression referring to death, and it has both inspired and been replaced by the more trendy expression, 'I'm archives.'"<sup>2</sup>

Today, according to a recent article in the "Sunday Styles" section of the *New York Times*, the trendy have taken their slang to an even higher level of sophistication. They are now studiously avoiding being associated with mundane activities such as "hosting" or "selecting," and are instead opting to engage in the more up-to-the-minute and stylish activity of "curating." The Oxford dictionary defines the standard meaning of "to curate" as "to look after and preserve." However, this sense of the word has been supplemented with a variety of non-traditional uses. The *Times* reported that "The Tipping Point, a store in Houston that calls itself a sneaker lifestyle shop, does not just sell a collection of differently colored rubber soles....No, its Web site declares, the store 'curates' its merchandise." Similarly, "Etsy, the shopping Web site devoted to

handmade and vintage goods, routinely brings in shelter magazine editors, fashion designers and design bloggers to serve as ‘guest curators.’” And “promoters at Piano’s, a nightclub on the Lower East Side [of Manhattan], have recently announced on their Web site that they will ‘curate a night of Curious burlesque.” Now if all of your competitors are “curating” merchandise, you do not want to be known as someone who merely “buys and sells” and, similarly, if all your rival nightclub promoters are “curating” parties, why in the world would you want to be left to be merely “hosting” one?<sup>3</sup>

In 1995, I was simply astounded at how change in popular jargon was so closely tracking a controversial definitional change in more esoteric circles. You’ll remember that one of the results of the task force was to loosen the definition of archival practice and extend some of its core concepts to define the practice of collecting and preserving digital information.<sup>4</sup> This definitional extension has now largely been accepted and even superseded, but at the time of its formulation, it was met with howls of protest from purists who felt that the task force was demeaning the value of true archival work by describing work on the ephemera of bits and bytes in the same terms. Find your own word, they said.<sup>5</sup> And today here we go again as the popular culture is closely tracking a more esoteric extension of the meaning of the term “curation” from museum practice to the definition of how effectively to manage and preserve floods of digital data produced by sensors of various kinds including telescopes, gene sequencers, and book scanners.<sup>6</sup>

What, if anything, do these various semantic extensions say about the value today of special collections, whether in artifactual or digital form? I will return to this specific question at the end of this paper. In the meantime, I want to explore some ideas about how best to construct the value proposition justifying investment in special collections, and about the areas of work that are likely to be most fruitful to advance scholarly communications.

## **The Definition of Special Collections**

“Special collections” is used in various senses for various purposes, sometimes referring simply to rare books and manuscript materials, and sometimes more generally to materials that are used as primary sources of evidence as opposed to secondary sources. In the recent working group report on *Special Collections in ARL Libraries*, “special collections” are defined “ecumenically” to include “any kind of vehicle for information and communication that lacks readily available

and standardized classification schemes, and any that is vulnerable to destruction or disappearance without special treatment.”<sup>7</sup> In this sense, special collections are those materials containing primary evidence for scholarship that require special treatment in their description or handling.

A value proposition is important because the costs of these special treatments can be quite substantial. At its most simplistic, the value proposition for special collections is that scholarship broadly across fields in the humanities, social sciences, and the sciences just cannot proceed without corollary investment in the acquisitions and carrying costs of the primary-source evidence needed to sustain and advance those scholarly fields. But how can or should a particular institution justify particular investments in particular kinds of collections? Tomes have been written on this more specific question. Institutional missions, areas of special expertise, previous investment in particular areas of scholarship, growth trajectories in new areas, and special opportunities presented by relationships with donors and private collectors are all among the factors that play a role in particular value propositions.<sup>8</sup> It is undoubtedly the complex nature of the interaction of these factors that accounts for the wide and rich variation among research libraries and archives in the kinds and level of their investment in special collections.

Added to the complex factors we know to be at work, the overall environment for scholarly communications has changed in startling ways and with these changes has emerged a new kind of conventional wisdom about special collections. Over the last 15 years there have been substantial not-for-profit and commercial investments in the electronic availability of back- and front-lists of journals and books that are of interest to scholars. What JSTOR, Project Muse, Elsevier and Wiley (among others) accomplished in the '90s for journals surely has many parallels to what Amazon, Google, and the Internet Archive (among others) have accomplished in the first decade of the new century for books. However, the massive Google books digitization project stands as a buoy marking the sea change that has occurred. As a way of taking account of these changes in the special collections arena, the conventional wisdom is to say that because books and serials are now more commonly available to wide audiences in the form of online networked information, what now makes libraries distinctive is not their book and serials holdings but their special collections.<sup>9</sup> Building on this conventional wisdom, it seems to follow logically that the value proposition for institutional investment in special

collections is that such investment is worthy because it will enhance the distinctiveness of the institution.

## A Critique of the Conventional Wisdom

This conventional wisdom about the distinctiveness of special collections compared to the commonness of book and serial collections certainly provides a useful heuristic and helps focus much needed attention on the requirements for building special collections into more useful scholarly resources. However, there are a variety of dangerous traps in the logic about common and special collections. First, system-wide analyses of research library holdings have suggested that books and serials that are being digitized are not so commonly

...there are a variety of dangerous traps in the logic about common and special collections.

held in libraries as one might have expected.<sup>10</sup>

With the lack of overlap, libraries cannot readily assume that their physical copies are represented in the common online collections, are held physically elsewhere, and thus can be readily

discarded. Instead, the digitization process may be accelerating the process of converting books and serials from circulating collections to collections of artifacts that need special treatment.<sup>11</sup> The most logical special treatment is not simply moving these artifacts into off-campus shelving but into more deeply rationalized and cost-effective shared shelving. In September 2009, both the Council on Library and Information Resources (CLIR) and the University of California received grants from the Mellon Foundation for separate but complementary, multi-institutional efforts to define the terms of and conditions needed to accelerate research library use of deeply shared storage facilities. This work builds on extensive previous work, especially by the Center for Research Libraries, the University of California, and OCLC's Programs and Research division.<sup>12</sup>

A second concern about the distinction between common and special collections is whether common collections that move online still require careful metadata treatments. Google, Amazon, JSTOR, and others with large aggregations of books and serials now provide access to inverted indexes, frequency analyses, and certain kinds of dynamically computed metadata such as a list of older works cited by a particular work and newer works that cite it. These search and discovery tools are proving to be a boon to scholarship. However, moving book and serial collections to the network has amplified,

rather than dissipated, other quality-control and metadata problems that are difficult to solve algorithmically and do require continuing special treatment. JSTOR has maintained the gold standard for descriptive metadata in its serials collections.<sup>13</sup> However, Geoffrey Nunberg has recently pointed to a variety of general errors in Google's book collection that are particularly troublesome for scholars who depend in their work on careful description of ordinary features such as series, edition, volume, and publication date.<sup>14</sup> In addition, the Council on Library and Information Resources will soon be releasing reports of extensive Mellon-funded studies by scholars in four different fields— linguistics; Latin American literature; history; and media history and cultural studies—that document vexing and ongoing quality-control problems in the book collections digitized by both Google and the Open Content Alliance.<sup>15</sup> Mellon also made a grant this summer to the University of Michigan for a systematic characterization of quality-control issues in the HathiTrust collections.

A third trap in the logic about the common and special collections lies in the largely unexplored area of what the Proposed Settlement Agreement for the Google book digitization project has called "non-consumptive research."<sup>16</sup> Joseph Esposito, Clifford Lynch, and others have often pointed out that the bulk of reading in the future will not be done by humans but by computers.<sup>17</sup> Non-consumptive research refers to such a kind of reading. Overall, our experience with non-consumptive research on texts is limited, especially in fields of the humanities outside of linguistics, but we have learned a good deal from the NORA, MONK, and SEASR projects at the University of Illinois at Urbana-Champaign. Teams of scholars led by John Unsworth, Martin Mueller, and others have found that computers are powerful readers when working on simple discovery tasks, but for advanced scholarly analysis, the machines are largely illiterate unless they are working on well-prepared and well-marked-up texts.<sup>18</sup> Different kinds of inquiries require different kinds of markup, often overlapping, and only some of the markup can be accomplished by algorithm given current technologies. Moreover, texts created by optical character recognition often need even further correction and preparation for sophisticated reading by machine. I assume that these various kinds of human intervention would be permitted on the texts stored in the non-consumptive research centers that the Google Settlement would establish. If not, much useful work could be done on public-domain materials even though the utility would be limited to special scholarly audiences in specific disciplines. In any case, the special

markup and error correction treatments required to make non-consumptive research, as opposed to simple search and discovery, truly useful to scholars puts the online collections of books and serials into a category that is far from common and more like the incarnation at the network level of the physical special collections that we know and love. Special collection skills and expertise are not unnecessary at the network level, they are simply operating in a different context.

The final trap I would mention lies in the suggestion that special collections are what give libraries and their home institutions their distinctiveness. Surely, special collections can be a source of pride, expertise, and excellence, and these

...taken to an extreme, the argument about institutional distinctiveness can also limit scholarly productivity by provoking the impulse to protect silo-like boundaries around collections....

qualities can motivate deep and useful investments. However, taken to an extreme, the argument about institutional distinctiveness can also limit scholarly productivity by provoking the impulse to protect silo-like boundaries around collections, thereby hindering the natural scholarly impulse to create and explore links

among related special collections across various holding institutions. Many have called for more openness within and connections across special collections,<sup>19</sup> but many barriers remain. I particularly invite library directors to

I particularly invite library directors to take a close look at the rights and permission statements that they have readers sign to use their special collections. Perhaps they will be as surprised as I was at the general, blanket, and highly restrictive claims their institutions make to usage rights over this material.

take a close look at the rights and permission statements that they have readers sign to use their special collections. Perhaps they will be as surprised as I was at the general, blanket, and highly restrictive claims their institutions make to usage rights over this material.<sup>20</sup>

I conclude from this brief critique of the conventional wisdom about the commonness of book and serial collections and the distinctiveness

of special collections that we need to refine our value proposition. The common versus distinctive opposition is simply too crude to get us very far. What is important about books and serials is that moving digital surrogates and newly produced works to the network level generates aggregations operating at a scale that advances existing lines of inquiry and opens new ones and makes scholars and students more productive, even when using individual works. These same criteria must form the heart of the value proposition for special collections.

Because special collections in the humanities, social sciences, and the sciences are full of primary-source materials, they are the fuel of scholarship in these areas. However, before making investments in them, libraries must answer: How would the investment advance existing lines of inquiry and open new ones? How would it make scholars and students more productive? Let me now offer for your consideration three potentially fruitful areas of activity for enhancing the value of special collections.

### Processing Special Collections

First, while there are many well-known, well-described, and heavily used special collections, the overwhelming problem that many research librarians have articulated in multiple conference papers and reports is the mountain of collections that remain unprocessed. Carol Mandel referred to the problem memorably as being like the “unwelcome white elephant” that eats you out of house and home.<sup>21</sup> CLIR’s Hidden Collections program is one small attempt at a solution. Perhaps more important has been the growing adoption of the “more product, less process” approaches that Mark Greene and Dennis Meissner have so effectively advocated.<sup>22</sup> Processing tools like the Archivists’ Toolkit and Archon have emerged and developers of both products are now working together to create a single unified product that consolidates the best features of

Deep knowledge of the collections is simply not sufficient for determining priorities for processing. Priorities must also be assessed against criteria of scholarly value

each and is better designed to operate and interoperate with related open source tools such as OLE, the Open Library Environment, and CollectionSpace, which is a museum-oriented system. We still lack the equivalent of a bibliographic utility for the detailed descriptions of special collections.<sup>23</sup> And because there is such

a mountain of materials to be processed, not as much attention has been focused as it should be on methods for efficiently determining priorities.

With Mellon support, a number of institutions have experimented with assessment tools to determine priorities for processing various types of collections. Although these tools now need to be accumulated, evaluated, and appropriately refined, libraries do need to use them more widely because it is amply clear from early experiments that they help focus library attention on the needs of scholars. Deep knowledge of the collections is simply not sufficient for determining priorities for processing. Priorities must also be assessed against

criteria of scholarly value, and for such assessments deep knowledge of the research and curricular priorities in various disciplines is also needed. In their forthcoming book from the Oxford University Press, Fran Blouin, Head of Michigan's Bentley Library, and Bill Rosenberg, a historian, analyze in detail the causes and consequences of the gulf in understanding that now exists between special collections librarians and scholars.<sup>24</sup> I urge you to read it when it is available. We urgently need creative solutions.

One way of gathering the deep scholarly knowledge needed to sort out priorities for special collection activity is to bring scholars and students directly into the special collections processing streams. Professor Jackie Goldsby is on the program of this meeting and will be speaking about pioneering efforts that she and her students have made with librarians and archivists at the University of Chicago and in the archives of various other institutions in and around Chicago. Mellon has funded similar initiatives at Columbia; Johns Hopkins; the University of California, Los Angeles; and the Huntington Library. All the results are not in yet, but what we do know is very promising, with benefits all around for the scholars and students, the library, and the university.

## Contribution Mechanisms

These programs illustrate one approach to the second fertile area of development for special collections to which I would draw attention: namely, finding efficient and productive ways to engage scholars and students in the development of special collections as scholarly resources. We have all heard about the Web 2.0 types of activities that try to draw readers in by adding tags or other forms of annotation to library records and surrogates. These are fascinating initiatives, but bringing scholars and students directly into the cataloging process is both more risky, and potentially more rewarding because of the deep engagement it can produce. Let me offer a few other examples to stimulate your thinking about how scholars and students could be productively engaged.

The Medici Granducal Archive in Florence, Italy, has a treasure trove of information about the Italian Renaissance that is almost entirely unprocessed. The Medici Archive Project (MAP), an organization based in New York, regularly provides residential research fellowships for visiting scholars, and hit on the idea in 1999 to develop a scholar-friendly data-entry system and require its fellows to spend a portion of their time cataloging the files they were



researching. One of the outstanding results of this project was the creation of a names identity database—a prosopography—that helps scholars sort out the identities of the formal and personal names that appear throughout the letters and other documents in the archives. After a decade of use, the data-entry system now needs to be upgraded, and MAP is using the occasion also to reconceive its fellowship programs. It will continue to have a small number of residential fellows, but is now planning for them to be of shorter duration so that it can also establish long-distance fellowships for individuals as well as a program for distance learning. For both of these new initiatives, MAP would digitize relevant files for the research or course topic but then still require the fellows and the students under the supervision of the course instructor to catalog at a distance these files as part of the interaction.

Another example is the work of Greg Crane, the classicist at Tufts, who established the Perseus database and has lately been hugely imaginative and productive in thinking about “What to do with a Million Books.” He and his research team have selected a corpus of books from the classical canon, worked with the University of Toronto and other libraries to ensure that these works find their way into the work flow of the mass digitization projects. His team then obtained library assistance and created a fully faceted, master bibliography of

...one solution to the metadata problems that are rampant in Google books might be to distribute the effort to self-organizing scholarly teams that care about specific parts of the corpus and will invest the necessary effort to correct and make it usable for scholarly purposes.

these selected works.<sup>25</sup> This initiative demonstrates that one solution to the metadata problems that are rampant in Google books might be to distribute the effort to self-organizing scholarly teams that care about specific parts of the corpus and will invest the necessary effort to correct and make it usable for scholarly purposes.

Crane and his team are also working with information specialists to engage other scholars and their students in developing and implementing the methods for applying linguistic markup to the corpus to facilitate machine analysis. Crane’s efforts seem to me to provide a model that could easily be emulated by other scholarly teams in other fields.

### **Connecting Collections**

This brings me to the third area of development I would ask you to consider: Can we develop new and reliable methods to link related special collections across institutions? We have been exploring this area at the Mellon Foundation

in several venues. Staff members of the archives at Boston University and Woodruff Library in Atlanta are together building a deeply integrated shared catalog of their holdings of the papers of Martin Luther King Jr. Into the project, they have drawn the scholarly editor of King's papers, whose team is contributing the vast knowledge it has accumulated about attribution, dating, provenance, and people. The project is also now considering how to draw in a third archives, the King Center in Atlanta.

The Integrating Digital Papyrology project based at Duke, with University Librarian Deborah Jakubs as one of the principal investigators, has gone a step beyond building a unified catalog by integrating three historically separate databases about essentially the same corpus of papyri: one containing bibliographic information; another containing images of the papyri; and the third containing transcriptions. Project staff are now in the process of adding an editorial overlay so that scholars can efficiently make new peer-reviewed entries into the database.

The Mellon Foundation also recently made a grant to a group of university presses, led by the Indiana University Press, all of which specialize in the publication of ethnomusicology. These presses have chosen to use as part of their publishing platform the database of Indiana's EVIADA project, a digital archives of ethnomusicological field video, so that primary source evidence can be closely linked to newly published monographs.

There are many other examples that I could offer from Mellon-funded programs, including the Roman de la Rose Digital Library led by Stephen Nichols, the Parker Library on the Web at Stanford and Cambridge, Electronic Enlightenment at Oxford, the Founding Fathers' papers at the University of Virginia, the Long Civil Rights Movement project at the University of North Carolina, and the Stalin archives at Yale. However, I hope I have said enough to convince you that a value proposition for special collections that is framed in terms of scholarly objectives is enormously attractive and opens a rich area for innovation and the pursuit of new lines of inquiry in a variety of scholarly fields.

\* \* \* \* \*

Now, in closing, let me return to a question that I raised at the beginning: What, if anything, do the various slang expressions about archives and curating that I mentioned at the beginning say about the value today of special collections, whether in artifactual or digital form? A flip answer would be to quote George Bernard Shaw, who once wrote that "people exaggerate the value

of things they haven't got....Everybody worships truth, purity, unselfishness for the same reason—because they have no experience of them."<sup>26</sup> Following Shaw, we could simply dismiss the slang as the inflated, self-important expressions of the unknowing. But we know Shaw to be wrong and so I commend to you the response of Laura Hotman, a senior curator at the New Museum of Contemporary Art. The author of the *Times* article on curating asked Ms. Hotman what she thought of the slang expression. "It doesn't really bother me," she said. "Actually, I'm hoping its popularity will spawn a reality show—maybe 'Top Curator.'"<sup>27</sup> Wouldn't that be fun!

*Acknowledgements: I would like to thank my Mellon Foundation colleagues, Helen Cullyer and Susanne Pichler, for their help at various stages in the composition of this paper. Paul Duguid also read a version of the paper and offered a number of helpful comments.*

- <sup>1</sup> Donald J. Waters, "The Implications of the Draft Report of the Task Force on Digital Archiving," in *Proceedings of the 127th Membership Meeting of the Association of Research Libraries* (Washington DC: ARL, 1995), <http://www.arl.org/resources/pubs/mmproceedings/127mmwaters.shtml>.
- <sup>2</sup> William Safire, "Kiduage," *New York Times Magazine*, October 8, 1995: 28, 30, <http://www.nytimes.com/1995/10/08/magazine/on-language-kiduage.html>.
- <sup>3</sup> Alex Williams, "On the Tips of Creative Tongues: Curate," *New York Times*, Sunday Styles section, October 4, 2009: 1, 8, <http://www.nytimes.com/2009/10/04/fashion/04curate.html>.
- <sup>4</sup> Donald J. Waters and John Garrett, eds., *Preserving Digital Information: Report of the Task Force on Archiving of Digital Information* (Washington DC and Mountain View CA: The Commission on Preservation and Access and the Research Libraries Group, 1996), <http://www.oclc.org/research/activities/past/rlg/digpresstudy/final-report.pdf>.
- <sup>5</sup> *Ibid.*, 46, n. 5.
- <sup>6</sup> See for example, Philip Lord et al., "From Data Deluge to Data Curation," in *Proceedings of the UK e-Science All Hands Meeting 2004* (Nottingham UK: UKOLN, 2004): 371–375, <http://www.allhands.org.uk/2004/proceedings/proceedings/proceedings.pdf>; Elizabeth Yakel, "Digital Curation," *OCLC Systems & Services* 23, no. 4 (2007): 335–340.
- <sup>7</sup> *Special Collections in ARL Libraries: A Discussion Report from the ARL Working Group on Special Collections* (Washington DC: ARL, 2009), <http://www.arl.org/bm~doc/scwg-report.pdf>.
- <sup>8</sup> See for example, Mary Lynn McCree, "Good Sense and Good Judgment: Defining Collections and Collecting" (1975), reprinted in Maygene F. Daniels and Timothy Walch, eds., *A Modern Archives Reader: Basic Readings on Archival Theory and Practice* (Washington DC: National Archives, 1984): 103–113; Linda J. Henry, "Collecting Policies of Special-Subject Repositories," *American Archivist* 43, no. 1 (1980): 57–63; William L. Joyce, "The Evolution of the Concept of Special Collections in American Research Libraries," *Rare Books and Manuscripts Librarianship* 3 (1988): 19–29; Michael T. Ryan, "Developing Special Collections in the '90s: A fin-de-siècle Perspective," *Journal of Academic Librarianship* 17 (1991): 288–293; Susan M. Allen, "Special Collections Outside the Ivory Tower," *Library Trends* 52, no. 1 (2003): 60–68; Frank Boles, *Selecting & Appraising Archives & Manuscripts* (Chicago: Society of American Archivists, 2005).
- <sup>9</sup> See for example, Beth M. Whittaker, "'Get It, Catalog It, Promote It': New Challenges to Providing Access to Special Collections," *RBM: A Journal of Rare Books, Manuscripts, and Cultural Heritage* 7, no. 2 (2006): 121; Rick Anderson, "Future-Proofing the Library: Strategies for Acquisitions, Cataloging, and Collection Development," *Serials Librarian* 55, no. 4 (2008): 563; Paul S. Koda, "The Past Is More than Prologue: Special Collections Assume Central Role in Historical Research and Redefine Research Library Collections," *Library Quarterly* 78, no. 4 (2008): 474; Sarah M. Pritchard, "Special Collections to the Fore," *Portal* 9, no. 2 (2009): 177; cf. Anthony Grafton, "Future Reading: Digitization and Its Discontents," *New Yorker*, November 5, 2007: 50–54.
- <sup>10</sup> Brian F. Lavoie et al., "Anatomy of Aggregate Collections: The Example of Google Print for Libraries," *D-Lib Magazine* 11, no. 9 (2005), <http://www.dlib.org/dlib/september05/lavoie/09lavoie.html>; Roger C. Schonfeld and Brian F. LaVoie, "Books without Boundaries: A Brief Tour of the System-Wide Print

- Book Collection," *Journal of Electronic Publishing* 9, no. 2 (2006), <http://dx.doi.org/10.3998/3336451.0009.208>; Brian Lavoie and Lorcan Dempsey, "Beyond 1923: Characteristics of Potentially In-Copyright Print Books in Library Collections," *D-Lib Magazine* 15, no. 11/12 (2009), <http://www.dlib.org/dlib/november09/lavoie/11lavoie.html>.
- <sup>11</sup> Stephen G. Nichols and Abby Smith, *The Evidence in Hand: Report of the Task Force on the Artifact in Library Collections* (Washington DC: Council on Library and Information Resources, 2001), <http://www.clir.org/pubs/reports/pub103/contents.html>.
- <sup>12</sup> Bernard F. Reilly Jr., *Developing Print Repositories: Models for Shared Preservation and Access* (Washington DC: Council on Library and Information Resources, 2003), <http://www.clir.org/pubs/reports/pub117/pub117.pdf>; Brian E. C. Schottlaender, "You Say You Want an Evolution...: The Emerging UC Libraries Shared Collection Concept," *Library Collections, Acquisitions, and Technical Services* 28, no. 1 (2004): 13–24; Lizanne Payne, *Library Storage Facilities and the Future of Print Collections in North America* (Dublin OH: OCLC Programs and Research, 2007), <http://www.oclc.org/programs/publications/reports/2007-01.pdf>; Constance Malpas, *Shared Print Policy Review Report* (Dublin OH: OCLC Research, 2009), <http://www.oclc.org/research/publications/library/2009/2009-03.pdf>.
- <sup>13</sup> Roger C. Schonfeld, *JSTOR: A History* (Princeton: Princeton University Press, 2003):156–160; Roger C. Schonfeld and Ross Housewright, *What to Withdraw? Print Collections Management in the Wake of Digitization* (New York: Ithaka S+R, 2009): 9–11, <http://www.ithaka.org/ithaka-s-r/research/what-to-withdraw>.
- <sup>14</sup> Geoffrey Nunberg, "Google's Book Search: A Disaster for Scholars," *Chronicle Review*, August 31, 2009, <http://chronicle.com/article/Googles-Book-Search-A/48245/>. See also Paul Duguid, "Inheritance and Loss? A Brief Survey of Google Books," *First Monday* 12, no. 8 (2007), <http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/1972/1847>; Grafton 2007, 52–53; and Robert Darnton, "The Library in the New Age," *New York Review of Books*, June 12, 2008: 72–80.
- <sup>15</sup> *Ghostlier Demarcations: Large-Scale Digitization Projects and Their Utility for Contemporary Humanities Scholarship* (Washington DC: Council on Library and Information Resources, forthcoming in 2010).
- <sup>16</sup> Proposed Settlement Agreement, *Authors Guild Inc., Association of American Publishers Inc., et al. v. Google Inc.*, case no. 05 CV 8136-JES, US District Court, Southern District of New York, filed Oct. 28, 2008: 11–12, 79–83, [http://www.googlebooksettlement.com/r/view\\_settlement\\_agreement](http://www.googlebooksettlement.com/r/view_settlement_agreement).
- <sup>17</sup> Joseph Esposito, "The Processed Book," *First Monday* 8, no. 3 (March 2003), <http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/1038/959>; Clifford A. Lynch, "Open Computation: Beyond Human-Reader-Centric Views of Scholarly Literatures," in Neil Jacobs, ed., *Open Access: Key Strategic, Technical, and Economic Aspects* (Oxford: Chandos Publishing, 2006): 185–193, <http://www.cni.org/staff/cliffpubs/OpenComputation.pdf>. Also see for example, Gregory Crane, "What Do You Do With A Million Books?" *D-Lib Magazine* 12, no. 3 (2006), <http://www.dlib.org/dlib/march06/crane/03crane.html>.
- <sup>18</sup> See for example, Tanya Clement et al., "How Not to Read a Million Books," paper presented at the Seminar on the History of the Book, Rutgers University, New Brunswick NJ, March 5, 2009, <http://www3.isrl.illinois.edu/~unsworth/hownot2read.rutgers.html>; Brian L. Pytlík Zillig, "TEI Analytics: Converting Documents into a TEI Format for Cross-Collection Text Analysis," *Literary and Linguistic Computing* 24, no. 2 (2009): 187–192; see also Gregory Crane et al., "Classics in the Million Book Library," *Digital Humanities Quarterly* 3, no. 1 (2009), <http://www.digitalhumanities.org/dhq/vol/3/1/000034.html>.
- <sup>19</sup> See for example, Daniel Traister, "Is There a Future for Special Collections? And Should There Be?" *RBM: A Journal of Rare Books, Manuscripts, and Cultural Heritage* 1 (2000): 70–72; Koda 2008, 479–480; Diane Zorich et al., *Beyond the Silos of the LAMs: Collaboration Among Libraries, Archives, and Museums* (Dublin OH: OCLC Programs and Research, 2008), <http://www.oclc.org/programs/publications/reports/2008-05.pdf>.
- <sup>20</sup> For an exception, see Peter Hirtle, "Removing All Restrictions: Cornell's New Policy on Use of Public Domain Reproductions," *Research Library Issues*, no. 266 (Oct. 2009): 1–6, <http://www.arl.org/resources/pubs/rli/archive/rli266.shtml>.
- <sup>21</sup> Carol Mandel, "Hidden Collections: The Elephant in the Closet," *RBM: A Journal of Rare Books, Manuscripts, and Cultural Heritage* 5 (Fall 2004): 106.
- <sup>22</sup> Mark A. Greene and Dennis Meissner, "More Product, Less Process: Revamping Traditional Archival Processing," *American Archivist* 68, no. 2 (2005): 208–263.
- <sup>23</sup> Whittaker 2006, 107.
- <sup>24</sup> Francis X. Blouin Jr. and William Rosenberg, *Processing the Past: Contesting Authority in History and the Archives* (Oxford: Oxford University Press, forthcoming in 2010).
- <sup>25</sup> Crane 2006, 2009; Christopher Blackwell and Gregory Crane, "Conclusion: Cyberinfrastructure, the Scaife Digital Library and Classics in the Digital Age," *Digital Humanities Quarterly* 3, no. 1 (2009), <http://www.digitalhumanities.org/dhq/vol/003/1/000035.html>.
- <sup>26</sup> George Bernard Shaw, "The Man of Destiny," in *Plays: Pleasant and Unpleasant: The Second Volume Containing the Four Pleasant Plays* (Chicago: Herbert S. Stone and Co., 1905): 311.
- <sup>27</sup> Williams 2009, 8.

© 2009 Donald J. Waters



This article is licensed under a Creative Commons Attribution-Noncommercial-Share Alike 3.0 United States License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/3.0/us/>.

**To cite this article:** Donald J. Waters. "The Changing Role of Special Collections in Scholarly Communications." *Research Library Issues: A Bimonthly Report from ARL, CNI, and SPARC*, no. 267 (Dec. 2009): 30–42.

<http://www.arl.org/resources/pubs/rli/archive/rli267.shtml>.

# Research Library Issues

## A Bimonthly Report from ARL, CNI, and SPARC

*Research Library Issues* is a freely available, online-only publication released six times per year by the Association of Research Libraries  
21 Dupont Circle NW, Suite 800  
Washington DC 20036  
voice: 202-296-2296  
fax: 202-872-0884  
ISSN 1947-4911

Executive Director: Charles B. Lowry

Editor: G. Jaia Barrett

Assistant Editor: Kaylyn Groves

Designer: Kevin Osborn, Research & Design Ltd., Arlington, Virginia

© 2009 Association of Research Libraries

ARL policy is to grant blanket permission to reprint from *Research Library Issues* for educational use as long as full attribution is made. Exceptions to this policy may be noted for certain articles. This is in addition to the rights provided under sections 107 and 108 of the Copyright Act. For commercial use, a reprint request should be sent to ARL Publications [pubs@arl.org](mailto:pubs@arl.org).

Current and back issues are freely available on the ARL Web site.

The Web site is also where you may sign up for alerts of new releases of *Research Library Issues* and learn about submitting ideas for articles.

Visit <http://www.arl.org/resources/pubs/rli/>.