The Idea of Order:

Transforming Research Collections for 21st Century Scholarship

June 2010

Council on Library and Information Resources
Washington, D.C.

ISBN 978-1-932326-35-2
CLIR Publication No. 147
Published by:

Council on Library and Information Resources
1752 N Street, NW, Suite 800
Washington, DC 20036
Web site at http://www.clir.org

Additional copies are available for $25 each. Orders must be placed through CLIR's Web site.
This publication is also available online at http://www.clir.org/pubs/abstract/pub147abst.html.

The paper in this publication meets the minimum requirements of the American National Standard for Information Sciences—Permanence of Paper for Printed Library Materials ANSI Z39.48-1984.

Library of Congress Cataloging-in-Publication Data

The idea of order : transforming research collections for 21st century scholarship.
p. cm. -- (CLIR publication ; no. 147)
Includes bibliographical references and index.
ISBN 978-1-932326-35-2 (alk. paper)
1. Research libraries--Forecasting. 2. Research libraries--Collection development. 3. Libraries--Special collections--Electronic information resources. 4. Research libraries--Book collections--Conservation and restoration. 5. Library materials--Digitization. 6. Scholarly electronic publishing. 7. Digital libraries. 8. Information commons. 9. Libraries and scholars. 10. Libraries and electronic publishing. I. Council on Library and Information Resources. II. Title. III. Series.

Z675.R45I34 2010
027.7--dc22

2010022593

Contents

Note: Full results of commissioned research on scholarly use of digital texts created through large-scale scanning projects, which are summarized in "Ghostlier Demarcations," are available online at http://www.clir.org/pubs/abstract/pub147.html.

About the Authors

Paul N. Courant is university librarian and dean of libraries, Harold T. Shapiro Collegiate Professor of Public Policy, Arthur F. Thurnau Professor, and professor of economics and of information at the University of Michigan. He has also served as provost and executive vice president for academic affairs at Michigan. He has published on a broad range of topics in economics and public policy, including tax policy, state and local economic development, gender differences in pay, housing, radon and public health, relationships between economic growth and environmental policy, and university budgeting systems. More recently, he is studying the economics of universities, the economics of libraries and archives, and the changes in the system of scholarly communication that derive from new information technologies.

Charles Henry is president of the Council on Library and Information Resources. Before his appointment early in 2007, he was vice provost and university librarian at Rice University, where he was responsible for library services and programs, including the Digital Library Initiative and the Digital Media Center. Mr. Henry is also publisher of Rice University Press, which was reborn in 2006 as the nation's first all-digital university press.

Geneva Henry is executive director of Rice University's Center for Digital Scholarship (CDS). The CDS supports digital scholarship projects at Rice and is home to the Digital Media Center, which provides multimedia equipment, services, and training to the campus to support the creation of rich media projects and scholarship. In 2006, Ms. Henry was a distinguished fellow with the Digital Library Federation, working with the Abstract Services Framework Working Group to develop a framework of digital library services. From March 2002 through June 2005 she served as executive director for the Connexions project, helping shape and launch the project. Prior to joining Rice, Ms. Henry was a senior information technology architect and program manager with IBM, where she was involved in several complex systems programs for government agencies, universities, and museums worldwide.

Matthew "Buzzy" Nielsen is assistant director at the North Bend Public Library in Oregon. His first library job 15 years ago was at the Langlois Public Library in Oregon, which serves a community smaller than most people's high school graduating classes. Currently, his interests include rural library development, library economics, nonprofit law and policy, and promotion of open source software in governments and nonprofits.

Roger C. Schonfeld leads the research efforts at Ithaka S+R. His work has focused on shifts in faculty attitudes and research practices in an increasingly electronic environment; shifts in teaching practices and the future of

instruction; and the academic library in a digital environment, including the economics, preservation, and policy issues associated with the transition from print to electronic formats for scholarly literature and government documents. He has served on the NSF Blue Ribbon Task Force on Sustainable Digital Preservation and Access, and on the Western Regional Storage Trust's advisory committee. Previously, Roger was a research associate at The Andrew W. Mellon Foundation. There, he wrote *JSTOR: A History* (Princeton University Press, 2003), which examines business models for the shift to an online environment for scholarly texts by focusing on how JSTOR developed into a self-sustaining not-for-profit organization.

Kathlin Smith is director of communications at the Council on Library and Information Resources where, over the past decade, she has coordinated the publication of more than 80 reports. She also edits and writes for the bimonthly *CLIR Issues* newsletter.

Lisa Spiro directs Rice University's Digital Media Center, where she helps plan and manage digital projects, studies the impact of information technology on higher education, and oversees a multimedia lab. She created the Learning Science and Technology Repository, founded the Digital Research Tools (DiRT) wiki, and served as coprincipal investigator of the Travelers in the Middle East Archive. A Frye Leadership Institute fellow, Ms. Spiro has published and presented on a range of topics, including the history of reading and publishing in nineteenth-century America, the impact of electronic resources on humanities scholarship, and text encoding. Her blog, Digital Scholarship in the Humanities, is at http://digitalscholarship.wordpress.com/.

Acknowledgments

CLIR would like to thank Alice Bishop, Charles Clotfelter, Amy Friedlander, David Gift, Michael Ann Holly, Clifford Lynch, Michael McPherson, Carole Moore, Stephen Nichols, and James Williams for reviewing and providing comments on the essays in this volume; and Leslie Eagle, Linda Harteker, and Brian Leney for their superb help in manuscript preparation. We also thank James O'Donnell for serving as coprincipal investigator on the project to examine large-scale text digitization projects and their use in contemporary humanities scholarship, and Jonathan Culler, Helen Cullyer, Anthony Grafton, Rick Luce, Stephen Nichols, Eleonore Stump, Will Thomas, Nancy Vickers, and Steven Wheatley for contributing to the discussion. Finally, we thank The Andrew W. Mellon Foundation for supporting the evaluation and analysis of large-scale text digitization projects.

The Idea of Order

Charles Henry

"The general aim is to piece together a new 'epistemic self-portrait': that is, a fresh account of the capacities, processes, and activities, in virtue of which Man acquires an understanding of Nature, and Nature in turn becomes intelligible to Man." Thus, in 1972 Stephen Toulmin laid out his complex goal in *Human Understanding: The Collective Use and Evolution of Concepts*.

In writing his book, Toulmin broke with the long tradition of Cartesian reasoning that strove to explicate human understanding in terms that described a set of concepts that were universally shared, immutable, and abstracted from the world around us. Toulmin rejects this approach and grounds his exploration of how concepts are shared, generation to generation, and how they may mutate in the process, by studying the influences of sociohistorical processes and intellectual procedures on collective understanding; by comparing historical and cultural contexts of various positions; and by including recent discoveries in psychological and physiological research that have focused on how the brain acquires and retains knowledge. His emphasis on scientific concepts in this regard stands as a lesser-known alternative to Kuhn's *The Structure of Scientific Revolutions*, though Toulmin also explored concept formation and evolution in engineering and the applied sciences.

The first important aspect of Toulmin's work is its groundbreaking, meticulous analysis of human understanding by virtue of our historical and cultural contexts as well as by the further permutations afforded by the wiring of our brains. Toulmin argues that the tradition of Cartesian (and Platonic) abstraction does not adequately represent the way we think and understand in real-world circumstances. To grasp human understanding, we must root ourselves in situational circumstances such as the workplace and the academic discipline; we must also note the customs of communication and

the forums for idea exchange these settings entail. The medium of academic life, for example, as well as that of neurological networks inside our heads, are salient environments for the collective use and evolution of concepts.

The second important feature of Toulmin's work is the analog nature of his study. Because it was written in 1972, there is no mention of the Internet, or a computer, or the Web; likewise, there is no mention of the words *digital* or *software*. The means of concept transmission in Toulmin's world is slow by today's standards: there are numerous references to academic disciplines, and the most common timeframe posited is generational.

For this CLIR study, it is useful to disaggregate the academic culture in which Toulmin was writing nearly 40 years ago from his methodology. We can easily appreciate differences in the pace and procedures of higher education between then and now, but the medium—the networks, processes, customs, and historical influences—are worlds apart. In today's era of astonishingly high-speed communication and scale of digital resources and tools, is human understanding different? The answer may be that it is, and perhaps profoundly so.

The recent collection of essays, *Understanding Knowledge as a Commons: From Theory to Practice*, explores the implications of relocating from an analog to a digital commons for creating, sharing, and preserving knowledge. One of the salient features of the new digital environment is what editors Charlotte Hess and Elinor Ostrom term *hyperchange*—change that is rapid, exponential, discontinuous, and chaotic. Aspects of hyperchange include increasingly permeable boundaries between knowledge creators, publishers, and readers; more flexible iterations of the processes and products of scholarly communication; the rise of new methodologies; greater collaboration within and among disciplines; and a more porous flow of original research among undergraduates, graduates, and faculty.

The analog "commons" depicted by Toulmin was a bounded space. For all the complexity that Toulmin introduced, the concepts with which he dealt evolved in ways that were intuitively organized and generally rational. Today's digital commons, by contrast, is often a contested zone where bounded and unbounded impulses compete: intellectual property laws, copyright, and the commodification of information can struggle with open access, file sharing, social networks, and a much more free-form, nonhierarchical, even chaotic participation in the creation and distribution of knowledge. The unbounded features of the new digital knowledge commons have resulted in the reconceptualization of academic libraries and, by extension, of the modern university. As Nancy Kranich notes, "self-governance of these newly emerging commons will require definition of boundaries (which tend to be 'fuzzy'), design and enforcement of rules, extension of reciprocity, building of trust and social capital, and delineation of communication channels. With research resources diffused throughout the campus and beyond, their broad scope requires stewardship well beyond the boundaries of the

edifices or structures that defined them in the past" (106).

Although each of the three essays in this volume arrives at its conclusions from a different lens, they are thematically coherent. All three show that our collective relocation from an analog to a digital environment for knowledge access, preservation, and reconstitution is under way and inexorable: the future of libraries and universities is digital. The first essay, "Can a New Research Library be All-Digital?" explores the degree to which a new research library can eschew print. "On the Cost of Keeping a Book," the second essay, provides a persuasive argument that from the perspective of long-term storage, digital surrogates offer a considerable cost savings over print-based libraries. The third essay, "Ghostlier Demarcations," describes research conducted against large-scale digital text data sets.

None of these essays, however, can be considered celebratory. A new research library cannot presume to be completely reliant on digital resources. A hybrid model of electronic and print materials will need to be juggled and budgeted for the foreseeable future. While the costs associated with a digital library are conceivably less than those of the nineteenth- and twentieth-century models we have inherited, numerous policy issues need to be addressed if digital surrogates are to be trusted, preserved over time, and duly audited. Complex migration strategies are also necessary and will add cost. "Ghostlier Demarcations" is perhaps the most cautionary of these essays. While a greater reliance and dependency on digital resources is inevitable, the quality of the data and their organization and accessibility in service to teaching and scholarship are major concerns. Without the guiding voice of scholars, the tremendous effort now being devoted to digitizing our cultural heritage could in fact impede, not facilitate, future research. Taken together, these reports amplify an intuitive insight: our migration to the digital commons will succeed only with the assumption of greater responsibility for its management, design, and sustainability. We must constantly and consistently rethink and reinvent as we trade places with an ancient and once comfortable analog world.

This volume can be viewed as a logical sequel to previous CLIR publications. Among them are *Diffuse Libraries: Emergent Roles for The Research Library in the Digital Age* (2002); *Library as Place: Rethinking Roles, Rethinking Space* (2005); *Preservation in the Age of Large-Scale Digitization* (2008); and *Working Together or Apart: Promoting the Next Generation of Digital Scholarship* (2009). A fifth CLIR publication, *No Brief Candle: Reconceiving Research Libraries for the 21st Century* (2008), called for libraries and universities to demand change; to create new career paths that better conform to emerging needs and methodological shifts in research and pedagogy; to conceptualize a research library as a multi-institutional entity; and to develop instructional and delivery mechanisms that are based on what we know of human learning and discovery.

This volume builds upon observations and recommendations in *No Brief Candle*. It is the first of a series of studies that will focus on services, responsibilities, and functions that have shaped the

traditional research library. *The Idea of Order* addresses the fundamental topic of collections: as the prevalence of physical objects fades into an environment of digital assets, the idea of a library and the nature of research become transformed. *Digital* is not a substitute for or replication of *analog*; we are evolving toward a new concept of a library but have neither experience nor precedent to comprehensively articulate this transition.

The three essays in this volume provide a specific and detailed perspective of what needs to be considered in the way of digital resources, economic models, and scholarly productivity to inform that reconceptualization, putting new stakes in new ground, fixing our lights to portion an encompassing sea.

Works cited

Hess, Charlotte, and Elinor Ostrom, eds. 2007. *Understanding Knowledge as a Commons: From Theory to Practice*. Cambridge, MA: MIT Press.

Kranich, Nancy. 2007. Countering Enclosure: Reclaiming the Knowledge Commons. In *Understanding Knowledge as a Commons: From Theory to Practice,* ed. C. Hess and E. Ostrom, 85–122. Cambridge, MA: MIT Press.

Kuhn, Thomas Samuel. 1962. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

Toulmin, Stephen. 1972. *Human Understanding: The Collective Use and Evolution of Concepts.* Princeton, NJ: Princeton University Press.

Can a New Research Library Be All-Digital?

*Lisa Spiro and Geneva Henry**

1. Introduction

What is the future of the library in the digital age? This question has preoccupied librarians and futurists since at least the 1960s, when prognosticators began to debate the significance of electronic information for libraries. In 2002, the *Chronicle of Higher Education* reported that college administrators and trustees believed that "the book will soon be the information medium of the past, if it is not already" (Carlson 2002). More recently, a CNN article announced that the "stereotypical library is dying," as physical books play a diminished role and librarians instead emphasize multimedia and the library as a social space (Sutter 2009). Daniel Greenstein, vice provost for Academic Planning, Programs, and Coordination at the University of California's (UC) Office of the President, predicted that the future university library would have a small staff, outsource many of its operations, and have a small physical footprint focused on special collections and study areas (Kolowich 2009). Yet some reject the vision of the virtual library, insisting that books are central to the library's role, and that it would be foolish to supplant our print heritage with the ephemerality of digital information. They believe that libraries will continue to function as hybrid libraries, offering access to both print and digital resources as appropriate.

We can bring the debate about the future of the library into sharper focus by asking how we would create a library at a *new* research university. Starting fresh can mean returning to first principles and planning for the future unencumbered by much of the weight of the past. Do new libraries need stacks filled with legacy collections of

* We would like to thank CLIR for its generous support in funding this study; Cynthia Gillespie for research assistance; the interviewees for sharing their insights with us; and the CLIR reviewers.

books and back issues of journals? What infrastructure is necessary for libraries to support the university's teaching and research missions? What kind of services should a twenty-first-century research library provide? Established libraries face significant challenges in moving toward mostly digital collections—challenges such as how to handle their legacy print collections; inadequate physical facilities; and faculty members, students, and library staff who are invested in the current modes of operating. Yet a new library faces the challenges of providing the level of space, staffing, and collections needed to support the university's mission at the outset rather than of gradually growing into a research library, as have so many of today's great research libraries. By starting anew, libraries can re-imagine their core mission and, in a sense, instantiate the twenty-first-century library.

Of course, startup libraries, too, must work in the present environment, where publishers do not yet provide access to electronic versions of many books, e-book readers have not been widely adopted and do not yet have many of the capabilities desired by researchers, solutions to the preservation of both print and digital collections remain uncertain, and libraries have accumulated prestige and expectations associated with maintaining print collections. Nevertheless, starting fresh gives libraries more freedom in defining what infrastructures and services are needed in the present and future rather than being constrained by the past. As Constance Malpas comments, "We've built up infrastructures that are all bound up in print materials. The sheer physicality of our collections has led us to come up with resource and staffing models that are not a good fit for our digital environment. That's what's liberating and a little scary about founding a new research library" (Spiro 2009h).

How to establish a new research library[1] is not simply a theoretical question. Since 2000, several new private colleges have been established in the United States, along with at least one research university and one state university. More than 20 health science schools and branches have been founded or are in the planning stages. Internationally, a number of new academic institutions have been founded. These universities are establishing libraries for the digital age—libraries that typically offer access to more electronic resources than print and that occupy facilities focused as much on collaborative spaces as on collections.

This report begins by examining the feasibility of establishing an all-digital research library through an extensive literature review.[2] It then offers brief, preliminary case studies of how several academic institutions founded since 2000 approached creating a new library from scratch. By "all-digital library," we mean a library

[1] Although our report focuses on research libraries, which are libraries at institutions classified as research universities by the *Carnegie Classification of Institutions of Higher Education*, we also look more generally at academic libraries, which support institutions of higher education such as colleges and universities.

[2] For an extensive and frequently updated bibliography of works on the future of libraries that was used to produce this report, see http://www.zotero.org/lisaspiro/items/collection/143163.

that provides only online access to its collections of journals, reference works, and books. It may continue to use interlibrary loan to procure print resources not yet available digitally; hold manuscripts, books, and artifacts in its archives and special collections; and offer study and collaborative space as well as in-person information services. We synthesize recent research about the technical, cultural, economic, and policy issues facing an all-digital library, particularly the problems surrounding e-books. Although we recognize that the environment is changing so rapidly that making definitive claims is nearly impossible, we hope to offer a high-level, multifaceted view of some of the central issues facing twenty-first-century libraries. The obstacles we identify will likely be overcome in time, although some changes (such as cultural ones) will likely take longer than others. To understand how new libraries are approaching the challenges and opportunities that accompany the shift to digital information, we look in particular at their collection strategies, service models, physical facilities, and staffing, drawing information primarily from interviews with leaders of these libraries. For many new libraries, small staffs and tight budgets challenge leaders to be inventive in how they deliver services and manage collections. Examining libraries that were recently founded or are now in startup mode may help us see emerging trends for libraries more generally. As several interviewees told us, some startup libraries may be a few steps ahead of established libraries, but nearly all libraries are moving along the same path toward the future. The report concludes with recommended actions and identifies areas for future research.

2. The Feasibility of an All-Digital Library

What conditions must be met for a research library to offer all-digital collections? To serve the needs of scholars, there must first be a critical mass of research materials available electronically. To meet researchers' needs, electronic resources must also be of sufficient quality; that is, they must be authoritative and trustworthy. In addition, researchers need convenient, user-friendly ways to read works in digital formats, such as e-book readers or print on demand. Researchers also should be able to perform basic and emerging tasks, such as annotating, analyzing, organizing, and visualizing information as well as moving seamlessly among different texts. An underlying assumption of scholars who rely on digital materials is the availability of a trustworthy, robust infrastructure that provides core services such as power and network connectivity. It should be easy for the researcher to discover electronic resources and for librarians to identify, license, purchase, catalog, and provide access to them. Managing access to the collection should not be too costly or burdensome for libraries. Libraries and publishers must devise effective strategies to guarantee long-term access to digital resources as well as to preserve print resources.

For journals and reference materials, these conditions have largely been met, although preservation remains a concern. The major

sticking point is monographs, the main focus of our report. In 2001, Clifford Lynch outlined some of the key impediments to the adoption of e-books—challenges that for the most part still remain nine years later. These challenges include the absence of a "critical mass" of scholarly monographs; lack of ubiquitous, user-friendly devices for reading long-form works; rights concerns; and doubts about long-term access (Clifford Lynch 2001). In addition, both librarians and patrons may resist the transition to digital collections because of concerns about quality, usability, and long-term access (Henry and Smith 2010). In a 2009 survey by Stanford University's HighWire Press, librarians identified digital rights management (DRM) as the single most important factor that hinders e-book use for library patrons (Newman 2010). Yet over time, it is likely that solutions to these problems will be found and that libraries will eventually transition to providing mostly digital access to their collections, given the ongoing shift in the publishing industry toward electronic delivery of information, the trend of libraries devoting increasing portions of their collections budget to electronic resources, user demands for digital information, and the continuing development of supporting technologies such as e-readers.

2.1 The Rationale for the Digital Library

By examining how to create a new research library, we are reframing a common question: what is the future of the library? As information increasingly is available online, libraries are changing. Adam Corson-Finnerty, formerly director of library development and external affairs at the University of Pennsylvania Libraries, points to many indicators of the decline of libraries, including rising costs for journal subscriptions, falling gate counts and checkouts, and diminishing support from academic leadership and faculty in disciplines such as science. Corson-Finnerty recommends that libraries shift away from most physical spaces and collections, and focus instead on delivering library resources, services, and tools digitally and on redeploying librarians as "informationists" who participate on research teams and serve as data curators (Corson-Finnerty 2009, 5). David Lewis, dean of Indiana University-Purdue University Indianapolis Library, likewise acknowledges the challenges libraries face at the twilight of print, but he views this as a time of opportunity for libraries to serve their mission of "making knowledge available in communities and organizations" (Lewis 2007, 419). He articulates five strategies, including completing the migration from print to digital, retiring print collections and preserving them by developing shared print collections, renovating library space to serve learning and collaboration, providing embedded support for research and teaching, and focusing on curating rather than on purchasing content. Core to these strategies is reimagining libraries' traditional mission in a digital context, so that they focus on helping people find, manage, and use information and on preserving long-term access.

As libraries transition from accumulating large physical

collections to providing online access to research, they are confronted with core questions about their mission (Montgomery and Marion 2002, 97). For many years, the size of a collection defined a library's greatness (Courant 2006). Statistics issued by the Association of Research Libraries (ARL) ranked libraries according to size, that is, the number of books and journals in their collection and the size of their staff and expenditures. If a researcher's local library had a work that was needed, he or she could do research much faster. Yet in an era where almost any resource can be rapidly delivered digitally to a patron's computer or physically via interlibrary loan, collection size is no longer as much of a strategic advantage. Indeed, having large collections means that one has to expend resources to store, manage, and preserve them. Access, rather than ownership, becomes paramount. Thus Paul Courant suggests that a new library would focus more on service and preservation than collections, noting that "if we were starting from scratch today, with the technology we have and a blank slate of IP [intellectual property] law and practice, we would immediately invent archives and archivists. The rest of library function, I think, would be organized around services rather than collections per se, except for 'libraries as museums'" (Courant 2006). What makes a library unique today is not the size of its holdings but the quality and innovative nature of its services.

Although this report focuses on challenges confronting new libraries, established libraries are likewise facing pressure to shift to digital collections, especially given space limitations, budget constraints, and user needs (Grafton 2009a).

- *Space:* Space constraints motivate some libraries to consider transitioning to digital collections. For example, the Claremont Colleges Library faced a severe space crunch that prompted it to investigate collecting e-books rather than print volumes. Ultimately, the library decided to buy digital back files for journals and to move print journals, rather than monographs, to off-site storage since "e[b]ooks tend to be more expensive and not as available or usable as e-journals" (Spiro 2009a). The Welch Medical Library at Johns Hopkins moved most of its print collections to the Libraries Service Center, instead aiming "to bring the library and librarians to people where they work and when they need the information" (Oliver 2005). The library focuses on delivering information digitally, offering access to more than 5,000 electronic journals, 400 data bases, and 2,000 electronic books (Welch Medical Library 2009). As Michael Kronenfeld comments, "The space is too valuable to waste with acres and acres of bound science" (Spiro 2009m).
- *Economics:* Another reason that libraries may turn to digital collections is to save money, particularly on journals. Today most libraries are "hybrid libraries," offering access to both print and electronic resources. Libraries have had to be hybrid, because not all resources were available electronically and not all readers were able to access information online. But as e-journals reach the tipping point, libraries are asking how long they should maintain

access to both print and electronic versions of the same journals. Providing access to both formats increases expense, consumes more staff time, and takes up more space. Schonfeld and Fenton point to economic evidence indicating that libraries can save money by moving to electronic journals, but that maintaining the current hybrid system eliminates these savings and may indeed lead to greater expense (Schonfeld and Fenton 2005). Life-cycling studies have shown that the long-term costs of maintaining print collections will likely outweigh the initial acquisition cost (Courant and Nielsen 2010). In contrast, while the initial costs of acquiring electronic content may be higher, the long-term costs will be lower (Spiro 2009h; Courant and Nielsen 2010). Although university administrators understand why libraries hold both print and electronic collections, library leaders will increasingly need to justify "why it is important to continue supporting so-called legacy collections and services as vigorously as electronic ones" (Beverly P. Lynch et al. 2007, 226).

- *User needs and services*: A fundamental objective of libraries is to support their patrons, so a key reason for providing access to electronic resources is to meet patron demand and enrich services. As more information becomes available digitally and on the network, libraries can provide new services to researchers, helping them search across distributed collections, mine vast repositories of data, mashup and manipulate information, and create custom collections (Dempsey 2006). Researchers appreciate the speed and convenience of electronic access. Further, they can perform full-text searches to discover information, save time by copying and pasting quotations, and store their research collections on a hard drive rather than in multiple file cabinets. Usage of electronic resources is high and increasing, and patrons are requesting access to even more (Martell 2007).

2.2 Obstacles to the All-Digital Library

Even though many libraries see strategic advantages in transitioning to digital collections, significant obstacles stand in their way. These obstacles include the lack of a critical mass of electronic resources; challenges faced by university presses and small publishers in transitioning to e-books; lack of appropriate means for reading long-form works; limitations imposed by DRM; resistance by researchers and librarians; difficulty integrating e-books into library workflows; concerns about long-term access and preservation; and economic considerations. The challenges are more acute for books than for electronic journals and reference works, which are already approaching the tipping point in adoption by users and libraries. While we have not treated each format separately to identify the obstacles in a more targeted manner, we do recognize that each format presents different challenges.

2.2.1 Lack of a critical mass of electronic resources

Before they shift to all-digital delivery of information, libraries must ensure that enough materials are available to support research and teaching. If libraries needed to provide access only to journals and reference collections, they probably could go all digital, since most of this content is already available online. As David Lewis observes, "There are three types of material to be considered as we look at the migration from print to electronic formats: reference works, journals, and books. The migration is nearly complete for the first two and is just beginning for the third" (Lewis 2007, 420). According to a 2007 study, "approximately 60% of the universe of some 20,000 active peer-reviewed journals is available in electronic form" (Johnson and Luther 2007). Moreover, 70 percent of libraries' subscriptions in 2006 were to e-journals or e-journals plus print. Although we are in an "extended transition zone" from print to digital, Johnson and Luther predict that over the next 5 to 10 years printed journals will decline, and ultimately 95 percent of journals will likely be electronic.[3] Several academic libraries that have opened since 2000, including the University of California, Merced (UC Merced) and California State University (CSU)–Channel Islands, have opted to subscribe almost exclusively to electronic journals, as have some established libraries such as Drexel. Libraries can see several benefits from subscribing to only electronic journals, including reduced costs, meeting researchers' preferences, freeing library space for other uses, and redirecting staff time to new priorities (Fenton and Schonfeld 2004; Roger Schonfeld 2007).

With monographs, the ability of a research library to rely on digital collections gets more complex. In 2003 Luther et al. concluded that "the potential of digitization has yet to be realized in the humanities, which tends to be monograph-dependent, because the e-book industry is immature and lacks economic and technological models necessary for large-scale adoption" (Luther et al. 2003). Although the e-book industry is maturing, those concerns still hold, since much content is not yet available electronically, business models are unsettled and multifarious, and universally satisfactory solutions for reading long-form scholarly works on a screen have not yet emerged. According to a 2008 study by Jason Price and John McDonald of Claremont Colleges, not enough recently published books are available electronically through aggregators for an academic library to pursue "paperless acquisition" (Price and McDonald 2008). Price and McDonald compared purchases of print books made by five academic libraries in 2006 and 2007 with the catalogs of four major aggregators of e-books for libraries (ebrary, NetLibrary, EBookLibrary, and MyiLibrary).[4] They found that around 70 percent of the libraries'

[3] ARL has commissioned a study from Luther and October Ivins to investigate whether libraries might assist "at-risk" peer-reviewed journals that are not yet electronic or lack an electronic subscription model. See Howard 2009b.

[4] Price and McDonald emphasize that their study focused on aggregators, not on the entire e-book universe. Although most publishers make their e-books available through aggregators, not every work is included. According to a 2008 survey of libraries, almost 70 percent of their total e-book spending was with aggregators (Primary Research Group 2008).

2006 and 2007 print acquisitions, including some of the most important books for researchers, were not available through these aggregators. According to their preliminary analysis, there is a mismatch between the content that some publishers (such as Routledge and Oxford University Press [OUP]) make available through e-book aggregators and what libraries typically purchase. For example, e-book aggregators included 400 OUP books that Claremont purchased during 2006–07 and provided access to an additional 1,555 books that they did not purchase, but did not include 150 books that they did purchase, which represented 37 percent of their total OUP acquisition (Spiro 2009a).[5] In some disciplines (art, music, romance literatures), more than 80 percent of library purchases are not available electronically, while in other disciplines (economics) nearly half (47 percent) are available as e-books. Price and McDonald conclude, "It looks like it's not yet an option to rely only on e-books for your monograph collection. … We found that the universe of what's available electronically right now is different from print" (Spiro 2009a). Their findings concur with a 2006 Joint Information Systems Committee (JISC) report that concluded that "the availability of 'core reading list' material (that is, those monographs and textbooks that are central to most academic programmes, and which are intended to be widely and intensively used by students) has been slow to develop. There are a number of reasons for this, of which the most important is probably that publishers have been reluctant to make available fully electronic versions of popular texts for fear of losing hard copy sales" (Higher Education Consultancy Group 2006). Many publishers have not yet figured out an appropriate business model for e-books and want them to supplement, rather than replace, print collections (Spiro 2009a).[6]

Nonetheless, e-books will likely play a significant role in research libraries in the next five years. In 2008, Carol Tenopir declared that "E-books Arrive" at libraries, pointing to data indicating high usage for electronic reference books and textbooks (Tenopir 2008). Mark Nelson, a fellow at the EDUCAUSE Center for Applied Research (ECAR), claims that technical obstacles and intellectual property concerns will likely be resolved within the next five years, and that cultural obstacles will recede as new, digitally savvy students enter college (Nelson 2008). Likewise, publishing consultant Anne Orens suggests that we are heading toward a tipping point for e-books, given the wide adoption of electronic journals, the growth in e-book revenues, the emergence of devices such as smart phones and e-book readers, the shift to a digital culture, print on demand, and Google Books (Orens 2009). In making the transition from print to electronic journals, libraries have already confronted and in large part resolved problems such as network connectivity and reliability of access (Prabha 2007).

[5] At the time of the study, Oxford produced electronic versions of their books, but did not make them all available through aggregators. In May 2008, OUP announced a partnership with Ingram Digital to make more titles available through MyiLibrary.

[6] To drive customers to print, several publishers are releasing e-books several months after print, much as they publish softcover books after hardcover editions.

In three to five years, Price and McDonald anticipate more e-books being available, whether through aggregators or directly from the publishers. Even in the past year, e-books seem to be catching on. According to the American Association of Publishers, in 2009 e-book sales were up 176.6 percent to $313 million, exceeding sales of audio books (American Association of Publishers 2010).[7] In the first quarter of 2002, e-book revenues were $1,556,499; in the first quarter of 2009, they were $25,800,000 (International Digital Publishing Forum 2009). Almost 90 percent of academic publishers have experienced growth in e-book sales, which now account for nearly 10 percent of book sales for the publishers responding to an Association of Learned Professional and Scholarly Publishers survey (Neilan 2010).

Libraries are among those purchasing e-books. A 2008 study by Primary Research Group found that library spending on e-books increased 36 percent between 2006 and 2007, although there was only a 13.6 percent increase in 2008 (Primary Research Group 2008). If the Google Books settlement is approved, libraries will have online access to millions of out-of-print, in-copyright books (Grafton 2009b).[8] New academic libraries may be at a disadvantage because they have not accumulated large legacy collections, but the Google Books settlement may partly undo that disadvantage and give smaller, newer libraries, as well as established libraries, access to a digital collection larger than the print holdings of many academic libraries.

There is a sense of inevitability about the transition to primarily digital collections reflected in many of the interviews we conducted and publications we examined. According to Bruce Miller, director of the library at UC Merced, the shift to digital has happened faster than he ever anticipated, as demonstrated by the high demand for e-books at his library (Spiro 2009i). By 2020, it is likely that most published resources will be in a digital format and that e-book readers and other reading interfaces will have evolved to the point that they are common and fully functional. According to a study cited by Lynn Brindley of the British Library, "by the year 2020, 40% of UK research monographs will be available in electronic format only, while a further 50% will be produced in both print and digital" (British Library 2005). While acknowledging that print will remain an important part of its collection, the British Library clearly recognizes the shift to digital information that is currently under way. Like other libraries,

[7] But as Sara Lloyd (2009) points out, e-book sales totaled "just 0.6% of overall book sales in 2008," so the steep percentage of growth in sales may be misleading.

[8] Google has reportedly scanned 12 million books. The Google Books settlement would allow Google to provide full-text access to in-copyright books that are no longer commercially available, about 70 percent of its total scanned books. Public libraries and not-for-profit higher education institutions could provide access to the full text of out-of-print but in-copyright works through a free public access service (PAS) terminal. (Associates colleges could provide one terminal per 4,000 students, while other higher education institutions could provide one per 10,000 students). Libraries could also subscribe to the Institutional Subscription Database (Band 2009). Library organizations have articulated concerns about the settlement, including privacy, equity of access to information, competition, and subscription pricing (ARL, American Library Association, and Association of College and Research Libraries 2009).

it is developing a digital strategy that includes providing researchers with access to a critical mass of digital content, connecting researchers to collections held by the British Library and elsewhere, building the digital infrastructure, and preserving both print and digital collections (British Library 2008).

Even if 90 percent of all research materials become available digitally, libraries will likely still need to maintain at least a small collection of print materials, as well as provide access to materials housed in shared print repositories. Not every book is well suited for an electronic format. Art and design books, artists' books, children's books, and even programming manuals still seem to work better in tangible, physical formats. Dee Magnoni of the Olin College of Engineering makes a persuasive case that "print is critical" (Spiro 2009o). For instance, Olin has an extensive collection of art and design books that simply would not work as electronic books (including one called *Spoon* that features a cover made out of metal and shaped like a spoon). As Magnoni says, these physical books unlock students' creativity: "Students come in and sit down, browse, pull them out, open them up. Their imaginations can go" (Spiro 2009o). Magnoni was surprised to find that students prefer print versions of computer and programming manuals, since they want to avoid switching between different windows as they program and consult the manual. Perhaps e-book readers will evolve to a point where they can simulate the experience of interacting with even the most elaborately designed physical book, but for now print continues to play an important role. Likewise, researchers will continue to interact with physical objects at archives and special collections, and historians of the book and other scholars will need access to original materials so that they can examine their physical characteristics as well as unique features such as marginalia.

2.2.2 Challenges faced by university presses and small publishers in transitioning to e-books

One reason why so few scholarly e-books are available is that scholarly publishers, particularly university presses, are finding it difficult to adapt to electronic publishing. Some university presses already have programs to deliver new or out-of-print books digitally; others would like to publish electronically, but face constraints such as inadequate funding, time, staff, and technological infrastructure. University presses have long-established workflows based on the print model. Changing these workflows would require presses, which typically have a tiny staff and operate on a shoestring, to commit significant time and resources to develop a new workflow and business model (Spiro 2009e). Given how leanly staffed many presses are, it is difficult for them to determine what electronic delivery platform to select and how to reconfigure their publishing processes at the same time that they are trying to publish print books under contract. Although university presses recognize that digital delivery is growing in significance, most of their revenues derive from selling print books, and they cannot abandon their core business. Making the shift

to publishing e-journals was easier (although not all small publishers have completed the transition) because presses could depend on a subscription income, there was a clear market, and the content was modular. Although presses can find savings in not having to pay for printing and warehousing books, they spend about 70 percent of the cost of a book before it is even printed (Spiro 2009e). Further, the business model for e-books may have unintended consequences. Where most journal income comes from library subscriptions, book publishers' revenue comes from a variety of sources. For many university presses, more revenue (though it varies for different publishers) comes from sales of books that are assigned for courses than from sales to libraries (Spiro 2009e). Publishers are concerned that making e-books widely available through the library may significantly hurt this important revenue stream, and libraries are unlikely to be able to pay enough more for e-books to make up the difference (Spiro 2009e). Another potential issue is the willingness of faculty to publish electronically, which is itself determined by tenure and promotion policies and the reputation of the journal or publisher (Harley et al. 2010). In some disciplines, digital publication is regarded as less credible than print. If faculty do not support electronic publishing, then university presses will be less likely to take the risk of moving to new publication models.

Some university presses are beginning to explore new publishing models and are moving into e-book publishing. For instance, the university presses at New York University, Temple, the University of Pennsylvania, and Rutgers recently won a grant from The Andrew W. Mellon Foundation to study "the feasibility of developing a consortium of university presses to deliver e-books to libraries on a shared platform" (NYU Press 2009). In 2006, Rice University Press was reborn as an all-digital press, using the infrastructure developed for the Connexions Project, which can produce electronic textbooks that can also be easily published as print-on-demand books (Buckman 2006). According to a recent survey by the Association of Learned Professional and Scholarly Publishers, 62.3 percent of academic publishers are now producing e-books, whereas only 17 percent do not plan to start publishing e-books in the near future (Neilan 2010). A recent ITHAKA report notes that presses may ease the transition to digital publishing by collaborating with libraries, bringing together publishers' skills in selecting, producing, and marketing scholarly content with libraries' expertise in describing, managing, and preserving information (Brown, Griffiths, and Rascoff 2007). Implementing such collaborations may be challenging, however, given funding constraints and differences in culture and organizational structures (Harley 2008).

2.2.3 Lack of appropriate means for reading long-form works

Modes of reading vary. Whereas researchers sometimes read to find a particular fact or quotation, at other times they want to immerse themselves in a text. Researchers interviewed by the Renaissance

English Knowledgebase/Professional Reading Environment (REKn/PReE) project have said that "the electronic is mostly used for information seeking and writing up of results, but pure reading is still handled in print" (Spiro 2009b). While reading on a computer screen may be appropriate for searching a text for a particular phrase, for many readers a desktop computer does not provide the kind of immersive reading environment that they prefer. E-book readers, which support a more ergonomic reading experience than reading from a computer screen, may provide a solution to the problem of reading long-form electronic texts. However, e-book readers have not yet been widely adopted, and early models lack many of the robust features that researchers require. A user study of e-book readers in graduate and undergraduate classrooms revealed three types of limitations in current-generation devices: (1) problems in how e-book readers operate as electronic devices, such as limitations in search, annotation, and touch-screen functionality; (2) functional limitations that are inherent in an e-book reader not being a print book, such as the inability to have more than one text open at a time, the lack of visual cues, and the difficulty of quickly navigating between texts; and (3) the lack of flexibility in the device, such as the inability to print or exchange books or to integrate them into other workflows (e-mail from Michael Furlough to Lisa Spiro, November 2, 2009). Researchers view the ability to annotate a book as important, but the annotation functions provided by current e-readers are not sufficient. For example, although the Kindle, currently the best-selling e-book reader, does provide a limited annotation feature and enables those annotations to be made available online through the annotator's Kindle account, a case study of the Kindle at Princeton showed that students found it difficult to take notes (Erick Schonfeld 2009; Princeton University 2010). Given the small size of the Kindle's screen and the inability of current e-ink technology to render color (only 16 shades of gray are displayed), it does not display photos, diagrams, and tables well, which are often essential to scientific works (Nicholson Baker 2009). Other devices that depend on e-ink technology, including the Sony Reader and the iRex, have similar limitations. Formatting often provides important contextual cues to the reader, but the formatting can be stripped or distorted in an e-book. Furthermore, citing an e-book is difficult, since no page number is given, just a "location range."

In addition to usability, affordability is an issue with e-book readers. Many students and faculty are unwilling or unable to spend approximately $250 for a Kindle or Sony Reader Touch Edition or approximately $500 for a Kindle DX. Although approximately 3 million e-book readers were sold in 2009 (most probably by Amazon), analysts also suggest that many people will not purchase an e-book reader unless it costs less than $100 (Lardinois 2010; Stone 2009b). If a device is required for readers to access library collections, libraries risk increasing the digital divide. According to librarian and e-book expert Sue Polanka, researchers like e-books because of the ease of access, but dislike them because of the difficulty of use. "The biggest

benefit of e-books is access, the ability for anyone to get it 24/7 … but the biggest drawback is, it's electronic. People generally want print and do not want to read on a computer. The devices are not compatible with all e-book formats and are too expensive for most users. They are not mainstream yet, so until the perfect device can be developed, we will always have the problem of great access but how do you read them?" (Spiro 2009g).

These problems with reading devices are likely temporary, as less expensive devices with more features are on the horizon and as alternatives to e-book readers such as tablet computers (including the Apple iPad), smart phones, handheld gaming consoles, and netbooks emerge (Bajarin 2009). The marketplace for e-book readers changes almost daily, as new devices, features, and publishing deals are announced. Manufacturers are working on color e-ink readers, which are projected to be available by the end of 2010 if not earlier (Ganapati 2009). In Japan, reading books on mobile phones is the norm. As of September 2009, 50 million iPhones and iPod Touches had been sold (Siegler 2009); many more people are potentially using the Kindle iPhone application than the Kindle device. Yet it remains to be seen whether a mobile phone is appropriate for scholarly reading, given the difficulty of annotating the text, comparing works, and scrutinizing illustrations on a small device. On the basis of a survey of students and staff at Open University and Cambridge University, Keren Mills concluded that it is premature for libraries to deliver e-books to mobile devices, since few cell phone users access e-books. "Most users," Mills noted, "are put off by the constraints of the technology, such as poor screen quality" (Mills 2009). Yet Mills also found that iPhone users are more likely to read e-books, which suggests that mobile phones do have some potential as reading devices.

It is unlikely that a single device will emerge as the dominant reading platform. Instead, researchers will likely use a diverse set of technologies to support reading based on their preferences and the tasks they need to perform. For instance, they may use smart phones for convenient, always-available access to reading materials, e-book readers for immersive reading, laptops or netbooks for quick searching and browsing, and desktops with large displays for more intensive visualization, analysis, or authoring. Hybrid devices such as the enTourage eDGe may offer the ergonomic benefits of an e-book reader alongside the information display and access capabilities of a netbook. Devices will likely provide new affordances for reading, such as the ability to annotate a text richly, record voice memos, have the book be read aloud, interact with a reading community online, and incorporate multimedia.[9] Ongoing research projects are examining how to create effective digital scholarly reading environments. For example, INKE (Implementing New Knowledge Environments)[10] will be developing ways to increase researchers' productivity in digital reading environments, such as by integrating research materials

[9] See, for instance, the Vook, which incorporates video as part of the text (Rich 2009a).
[10] http://www.inke.ca/.

from multiple sources in a single interface (which increases productivity six- to seven-fold) and designing more-effective interfaces (Spiro 2009b).

2.2.4 Limitations imposed by digital rights management

Even if inexpensive user-friendly devices become ubiquitous, researchers may be too frustrated to use e-books if publishers continue to impose overly restrictive DRM protocols. You can do a lot with a print book: photocopy or scan as many pages as you like, scrawl in the margins, highlight passages, bookmark pages, flip between it and other books, read it in the bathtub, give it to someone else, make art out of it, and so forth. Because of constraints imposed by some DRM regimes, readers of e-books may find that they can print only a limited number of pages, have to navigate awkwardly through the book, cannot take notes or bookmark pages, and cannot give the book to someone else. Moreover, DRM may deny them the ability to access digital content on a variety of devices, from cell phone to e-book reader to computer to print (although readers can access Kindle content on an iPhone, desktop or laptop computer, Kindle, and/or iPad).

DRM is not intrinsic to e-books or e-book readers: it is a choice that publishers make to "protect" their content. As Clifford Lynch argues, the rights of access that come with an e-book purchase result from a negotiation between publisher and customer, since "the e-book reader is fundamentally agnostic about the technological control of intellectual property" (Clifford Lynch 2001). Publishers adopt DRM because they worry about losing revenue as a result of unauthorized distribution of their work, but in imposing restrictive DRM they hold back adoption by the academic market, which values openness and usability. As Michael Furlough has commented, "I feel like the publishing industry is ready to experiment with devices, but they're where the music industry was five to eight years ago. They're not ready to be more open with content. … It's going to take more work to get this to work in the academic environment" (Spiro 2009d). Concerns about privacy and ownership of electronic content also impede adoption of e-books. Amazon recently deleted copies of George Orwell's e-books from customers' Kindles because of intellectual property issues, raising fears of censorship and corporate control of information (Stone 2009a). In an e-book world, does the library or the customer truly own content? Can an e-book be moved from device to device and migrated to new formats? If a company goes out of business, will the customer still have access to purchased content?

In addition to raising scholars' concerns about usability and openness, DRM makes it difficult for libraries to adopt e-books. In order both to lend and preserve books, libraries depend on the "first-sale" doctrine, which states that the purchasers of books may do what they want with them, including loaning or reselling them. But publishers typically license, rather than sell, e-books to libraries, so libraries are circumscribed in how they provide access to them. In the print world, a librarian can simply purchase a book; with e-books, each license must be reviewed carefully. Libraries are stuck with the

DRM regime that the aggregator negotiated with the publisher at the time of purchase (Spiro 2009a). Polanka notes that there is little consistency in licensing agreements: "You can work with 100 publishers and have 100 different agreements, 3 different aggregators and have 3 different agreements" (Spiro 2009g). Rather than having to call in a lawyer to review every license, libraries would prefer to have a standard license, which would streamline purchasing. In negotiating with vendors, libraries need to examine restrictions on loaning and using e-books, long-term access, the ability to move content onto multiple devices, the ability to update and migrate content, and privacy (Clifford Lynch 2001). DRM with textbooks poses additional challenges in adoption since textbooks are not licensed to institutions (Newman 2010). Given that many readers prefer not to read e-books on computer screens, can libraries make them available on e-book readers and/or via print on demand? Will multiple patrons be able to check out an e-book simultaneously? Can scholars perform sophisticated searches and textual analyses across large corpora of texts? Most publishers prohibit interlibrary loan for e-books, undermining a key goal and operational strategy for libraries, which are banding together in consortia to share resources and to lower costs (Ball 2009, 21). Concerned that restrictive licenses may "significantly reduce users' rights," including the right to loan e-books via interlibrary loan, reformat them for preservation or access by those with disabilities, and perform basic operations such as copying and pasting, the Canadian Association of Research Libraries (CARL) Copyright Committee Task Group on E-Books recommends that libraries create a set of principles for licensing e-books as well as a model license (CARL Copyright Committee Task Group on E-Books 2008, 14).

Recognizing the increasing demand for the ability to read books on e-book readers, some aggregators and publishers are exploring different possibilities for access (Lonsdale and Armstrong 2008). For instance, Springer offers favorable "interlibrary loan provisions," and other publishers are considering easing their restrictions as well (Li 2009). In response to concerns about how library content will be delivered to portable devices, Sony is working with libraries to enable patrons to download library e-books licensed from OverDrive onto the Sony Reader (Scott 2009). Likewise, the e-book library service EBL allows some of its e-books to be downloaded onto the Sony Reader, claiming that it is one of the first e-book providers to do so (Paulson 2008). A few libraries are experimenting with loaning Kindles preloaded with e-books, but it is not clear that this is legally permissible (Oder 2009a). Perhaps scholarly e-book publishers will follow the lead of some journal publishers and relax DRM in order to appeal to the scholarly market. Although most journal publishers impose access controls on their publications, they typically do not limit the ability to download or print a .pdf or .html version of the article, because researchers view this feature as crucial. As ScienceDirect has found, DRM imposes costs and creates support issues for publishers, so it gives researchers the ability to download articles from more than 2,000 e-journals (Biglione 2007).

2.2.5 Resistance by users

Perhaps the greatest obstacle to the transition to an all-digital library is resistance by users, particularly faculty. As Price and McDonald suggest, moving to all-digital collections would require "a cultural shift in information usage by our faculty," since a vocal minority still use research methods based on print resources and demand the same from their students (Spiro 2009a). As libraries change how they deliver collections and services, they need to take into account the needs and attitudes of faculty or risk losing support of vital stakeholders and undermining their research and teaching. Yet serving the faculty is tricky, since research practices vary by discipline. According to an ITHAKA faculty survey (Schonfeld and Housewright 2008), faculty in the humanities continue to rely on print resources, while those in the sciences are increasingly embracing the digital. The preferences of faculty in the social sciences lie in the middle, except for economists, who favor shifting to digital resources (Schonfeld and Housewright 2008, 13). Attitudes toward digital resources are dynamic, driven by the availability of tools and resources as well as by research methods. How researchers view the library is changing, too, as fewer faculty (particularly in the sciences) see the library as a gateway to information resources or begin their research on the library Web site (Schonfeld and Housewright 2008). Although faculty in the humanities have been slower to adopt digital technologies, the ITHAKA surveys demonstrate that they "are on basically the same trajectory as scientists, simply less far along" (Schonfeld and Housewright 2010, 34).

In response to changing attitudes toward electronic resources, libraries may feel pulled in multiple directions, depending on the disciplines of their users. For many humanities scholars, the library functions as a sort of lab where they find the resources and tools they need to do their research. In contrast, many scientists can do their research without setting foot into the library; they do most of their work using journals that are, for the most part, online (Grafton 2009a).[11] How should libraries serve both the digitally dependent and print-oriented users? Schonfeld and Housewright argue that libraries face a critical strategic dilemma:

> On one hand, the fields whose practices are most traditional also appear to contain the library's greatest supporters; therefore, if the library shapes its roles and activities based on what is currently most highly appreciated by faculty, it may lose a valuable opportunity to innovate and position itself as relevant in the future. On the other hand, if the library develops new and innovative roles and services that address unmet needs, becoming newly relevant and even essential to those scholars who have moved farthest away from it, in the near term it may lose the support of its most ardent supporter (Schonfeld and Housewright 2010, 14).

[11] Even in scientific disciplines there may be occasional need for resources that have not yet been digitized, such as conference proceedings or monographs.

Libraries risk becoming irrelevant if they ignore emerging needs and research practices, but they have also faced backlash for moving away from print collections. Yet as Schonfeld and Housewright also point out, all fields are heading toward increased reliance on digital content, albeit at different speeds. Libraries thus must determine how quickly to change and how to communicate those changes to users.

Even as they increasingly rely on digital resources, scholars articulate some significant concerns about the shift to a digital library. Many humanities scholars are enthusiastic about the ability to find so much information online, saying that remote access has been transformational. Yet there is also a sense of anxiety and loss, as researchers worry that the exclusive reliance on digital resources will change research for the worse and will lead to restrictions on, and the loss of, information (Clifford Lynch 2001). Although researchers are increasingly reliant on electronic access to current issues of journals, many worry about replacing print back files with electronic journals, concerned about the "reliability" and completeness of the electronic (Schottlaender et al. 2004). However, according to the ITHAKA S + R Faculty Survey 2009, slightly more than one-third of respondents felt that it would "*always* be *crucial* for *their own* college or university library to maintain print journal collections" (Schonfeld and Housewright 2010, 19). In many cases, humanities scholars' research depends on both print and electronic resources. Many scholars engage in "mixed practices," accessing their library's journal collections online but subscribing to print versions of core journals to keep current and build personal collections (e-mail from Sophia Kyzys Acord to Lisa Spiro, Feb. 15, 2010). Some are concerned that younger scholars know only how to search Google Scholar without hitting the library, thereby blithely overlooking important, and sometimes not-yet-digitized, resources and biting off disembodied chunks of information rather than digesting entire works that lay out a closely reasoned argument (Spiro 2009l). As the availability of information and the speed of research increase, these concerned scholars insist on "the need for a careful analytical research process" (Harley et al. 2010, 16). Still others worry about selective digitization, warning that many important resources—including those produced outside the United States and Western Europe—have not been digitized; they fear that scholarship may suffer if these resources are ignored (Harley et al. 2010). Given the sense of ephemerality and the perception of a lack of peer review associated with some forms of electronic publication, some faculty may not trust these resources as much as they do publications that are linked to a familiar and trusted imprimatur, which is perceived to be both more authoritative and more fixed and stable. Citation practices are still evolving, with many researchers preferring to cite the more authoritative print version than the digital version they actually consulted. The same e-book may be formatted differently for different delivery systems (for example, online, e-book readers, mobile phones, Kindle, Stanza), raising the question of which format should be regarded as authoritative. For

scholars across the humanities, a digital surrogate cannot always replace the original object. Such scholars need to work physically with the material or archival object, to see the edges of a manuscript and the color of the ink, to "hear the nuances of harmony, multi-channel directionality, and instrumentation not captured by MP3 compression," and to get an embodied sense of the physical dimensions of the original object (Spiro 2009l). Likewise, scholars interested in the history of reading do not want to lose evidence of reader responses, such as annotations made in the margins of books.

While undergraduate students typically do not approach research as comprehensively as faculty do, they, too, express a desire for continued access to print, as well as an enthusiasm for electronic resources. According to a 2006 study by OCLC, 89 percent of undergraduate students begin their research by conducting a Web search and 44 percent said they used an online database at least once a month; 39 percent borrowed a print book at least monthly (De Rosa et al. 2006). Students are likewise embracing e-books; at the University of Illinois, for example, 55.5 percent of undergraduate students reported that they had used e-books, mostly for study and research (Shelburne 2009). Yet when it comes to reading a complete book, many students still prefer print (Shelburne 2009). As e-book readers become more ubiquitous and more powerful, students may feel more comfortable reading longer works on a screen. However, pilot projects to test Kindles at Princeton University, Case Western Reserve University, and the University of Virginia's Darden School of Business found that students thought it was difficult to take notes and to navigate among multiple documents at once; they preferred the Kindle for leisurely, rather than scholarly, reading (Kolowich 2010). For now, it seems that students want access to both print and electronic books. Students have participated along with faculty in the protests against off-site storage at Syracuse and Ohio State. A 2004 survey of graduate (master's degree) students at San Jose State University showed that they preferred a hybrid library (Liu 2006). At UC Merced, one of the new academic libraries that we examine in this study, students responding to a 2009 survey called for more print books to be added to the collection (UC Merced Library 2009c). In the University of Illinois survey, 55 percent of undergraduate students predicted that they would be reading both print and electronic books in the future; 26.5 percent said only print, and 11.5 percent only electronic (Shelburne 2009).

Attitudes toward electronic books may change over the next 5 to 10 years, particularly as it becomes easier for researchers to accomplish their goals and develop new modes of doing research. Our experience with the adoption of e-journals shows that it takes time for people to embrace new information technologies. Tenopir et al. present an evolutionary model for scientific electronic journals: (1) the early phase (1990–1993), when e-journals were introduced; (2) the evolutionary phase (late 1990s–early 2000s), when most scientific articles were available electronically; and (3) the advanced phase (2002–present), when the full text of core journals' complete runs

and advanced features such as sophisticated search were provided (Tenopir et al. 2003). Scientists' use of e-journals dramatically increased across these evolutionary periods: "0.3, 38.8 and 79.5 percent of readings are from electronic format through early, evolving, and advanced phases" (Tenopir et al. 2003). Most recent studies show an increasing usage of and preference for electronic journals over print (Bar-Ilan and Fink 2005). Likewise, in the case of reference works, there seems to be broad consensus that the electronic format is preferable, given the ease of updating encyclopedias and dictionaries and searching across them. As Roger Schonfeld suggests, "Everyone feels that dictionaries, encyclopedias, etc., should be electronic. … This is an instance where the print format didn't work as well as electronic" (Spiro 2009k). Most researchers also agree that electronic resources will have an increasing impact on scholarship. According to the ITHAKA surveys, a growing number of faculty agree that "I will become increasingly dependent on electronic research resources in the future"—just over 60 percent in 2000, 70 percent in 2003, and about 75 percent in 2006. Likewise, the Center for Studies in Higher Education finds that "scholars are unanimously enthusiastic about the growing volume of research material located online and the development of powerful tools to process them," even as they worry about being flooded by information (Harley et al. 2010, 16).

There are important differences between e-journals and e-books. With short-form works such as articles, researchers can easily either read on the screen or print out the article. However, most people do not feel comfortable reading books on a computer screen, and printing out an entire book is too cumbersome, if it is even permitted. According to ebrary's 2007 *Global Faculty E-book Survey*, 658 of 829 (80 percent) respondents agreed that "when reading the whole book or extensive sections, print books are preferable" (ebrary 2007). Survey results indicated that advantages of print included ease of use, portability, and the ability to take notes and highlight passages. The advantages of e-books were accessibility at any time from any place and ease in searching and browsing. In describing what would make e-books more appropriate for their area, faculty gave top ranking to "greater breadth and depth of collection," "ability to download," and "less restrictions on printing and copying," suggesting that access, rights, and usability are the top issues. As Polanka observes, "In general, I still think most people want the print. They want to touch it, write on it, flip the pages. That has value for people. Culturally, that's what we're used to. It might take a while for people to get used to the idea of electronic books" (Spiro 2009g). In ITHAKA's *Faculty Survey 2009*, e-books ranked last in the list of electronic resources that faculty regarded as important to their research and teaching (Schonfeld and Housewright 2010). According to this survey, only 4 percent of respondents strongly agreed that "Within the next five years, the use of e-books will be so prevalent among faculty and students that it will not be necessary to maintain library collections of hard-copy books" (Schonfeld and Housewright 2010, 23). More respondents see e-books gaining importance in research and teaching. Although only

13 percent of respondents believe that e-books are important today, 31 percent think that they will be valuable in five years (Schonfeld and Housewright 2010).

Despite problems with usability and availability of scholarly content, some students and faculty are beginning to use e-books, although often as a complement to rather than replacement of print books. Several studies have found that as the usage of print books has declined, e-book usage has increased, particularly in fields such as computer science, business, and engineering (Cox 2008, 4-5). For instance, an analysis of e-book usage data at Oakland University shows users preferred to use electronic books in technical fields such as computer science (Slater 2009). Indeed, studies by the UK's CIBER research group suggest that "we have an e-book take-off," since researchers want the "condensed, distilled knowledge" offered by books and can find chunks of content in books more easily once they are made searchable online (Nicholas et al. 2008). Yet it seems that many people use e-books to find content, and print books to digest it. The JISC National E-books Observatory Project, for instance, found that students and staff prefer the electronic for "grazing and extracting information rather than for lengthy reading" (JISC 2009, 17). As Schonfeld and Housewright argue, "Neither faculty members nor librarians expect e-books to constitute a viable substitute for print books; they are more generally seen as complementary" (Schonfeld and Housewright 2008, 22).

Not only are physical books well-suited to long-form reading, but they and the libraries that house them play a profound cultural role—communicating, preserving, and symbolizing cultural memory. Even among some scientists who get most of their information online, there can be a sense of nostalgia about the traditional library building and the serendipity of browsing stacks (Spiro 2009l). Price and McDonald tell of a faculty member who asked the library to subscribe to Oxford Scholarship Online. When told that budget constraints meant that if the library bought online access it could no longer purchase the print versions of the books, the faculty member balked, maintaining that "no library can be without print editions of these core materials" (Spiro 2009a). Libraries are symbols of a continuity of past and present; they offer access to the cultural heritage and pledge to preserve it into the future. As libraries move into the digital future, they need to take into account anxieties about what may be lost: immediate access to print stacks, a tangible connection with the past, "an assurance of solidity and objectivity to a culture awash in postmodern skepticism" (Manoff 2001, 379). Even new library buildings often evoke the print tradition, such as by building grand reading rooms or making print collections more visible. As Steve Shorb of New York University Abu Dhabi (NYUAD) notes, working on several library construction projects has demonstrated to him that people expect a library to refer to the tradition of the library and offer some access to print books.

Possible resistance to the shift to all-digital collections may have been foreshadowed by protests against off-site storage at universities

such as Syracuse (Epstein 2009) and Ohio State (Howard 2009a). Some scholars insist that they need to be able to browse books on the shelves so that they can serendipitously discover related works. When Stanford proposed to shut down its East Asian library, move volumes to off-campus storage, and focus more on the digital delivery of information, a subcommittee of the Library Committee issued a report insisting on the continued importance of print collections and browsing. The report claims that "it will require at least two generations of faculty renewal—something like 50 years—before electronic media take precedence over paper support in some fields of inquiry" (C-LIB Subcommittee on Digital Information Technologies in the Research Library Environment at Stanford 2008, 11). While the C-LIB Subcommittee does not reject electronic resources, it insists that research libraries should continue to operate as "hybrid institutions," serving scholars' needs by providing access to both print and electronic resources (C-LIB Subcommittee on Digital Information Technologies in the Research Library Environment at Stanford 2008, 11).

Countering the trend toward off-site storage of print collections, the University of Chicago is offering a hybrid vision of the "library of the future" by erecting a new facility that will store collections on campus. The new Mansueto Library will bridge the print and digital worlds by featuring high-density shelving for 3.5 million additional print volumes, an automated storage-and-retrieval system, a grand reading room with seating for 150, a conservation laboratory, and a digital technology laboratory. The decision to build an on-campus facility was based on the recommendations of a faculty committee. A subsequent committee, the Provost's Task Force on the Library, explored ongoing changes in the library (Abbott et al. 2006). This task force conducted studies of library usage, such as a 2005 survey of undergraduate and graduate students that showed print and digital collections are "synergistic," as researchers who check out many books also use many electronic resources (quoted by Dixon 2005). While the committee found that electronic journals were supplanting print, it also concluded that print books continued to play an important role in research, particularly for approximately 500 to 1,000 "heavy users." The committee argued that the library's central mission should be to support research, and that the library could distinguish itself by resisting the trend toward off-site storage. In justifying building a high-density facility on campus, the library thus focused on researchers' need for "immediate on-campus access to integrated collections," the unlikelihood that in-copyright and unique materials will be available electronically, uncertainty about long-term access to digital collections, and the belief that "mass digitization leads users to collections" (Nadler 2008; University of Chicago Library 2008). Opposing the assumption that book collections are going obsolete, Judith Nadler, Chicago's university librarian, stated, "We believe, instead, that scholarship will thrive in an environment where print and electronic coexist, now and in the future" (Nadler 2008).

Even if many library patrons are comfortable with moving toward electronic resources, a few vocal faculty can block this

transition. Library leaders must possess political skills to convince faculty of the economic and intellectual advantages of shifting to electronic access, along with the appropriate skills for managing preservation and access risks (Spiro 2009k). Schonfeld points to a gap between practice and belief among faculty. Even though faculty rely upon electronic journals, they may be unwilling to give up print journals because they have little incentive do so. Subscription costs do not come out of their own budgets, and they may see real risks, such as uncertainty about long-term preservation, in moving toward electronic information (Spiro 2009k). Yet as Schonfeld also acknowledges, "Faculty attitudes will always lag behind faculty practices, but attitudes are steadily shifting, which is one of the reasons that libraries are ready to move in this direction" (Spiro 2009k). For example, most faculty no longer believe that their institution must assume local responsibility for preserving journals and reference works, as long as they are preserved somewhere (Schonfeld and Housewright, 2010). Attitudes toward electronic resources can change rapidly, as shown in the case of the economics discipline. Although the social sciences once relied on printed resources, more than 70 percent of economists surveyed by ITHAKA in 2006 favored canceling print subscriptions and subscribing only to electronic journals, and only 25 percent felt that the library should maintain its print journal collections for the long term.

As the 2006 ITHAKA report concludes, similar shifts may occur in other disciplines: "As new tools emerge and mature, however, the format which best supports scholarship may shift, and preferences and practices may shift to whichever format best facilitates scholarship" (Schonfeld and Housewright 2008, 14). As usability issues are worked out, quality and credibility concerns are resolved, and researchers take advantage of the convenience and analytical power of electronic access, much of their resistance will likely dissipate. Budget challenges may drive the shift to electronic resources, particularly as libraries have difficulty paying for both print and electronic journal subscriptions. Library directors answer not only to faculty but also to administrators. The administrators who provide library budgets may be reluctant to fund new facilities to house print collections and may question large expenditures to support both print and electronic formats. Library directors must consider not only the immediate expectations of faculty but also the long-term goals for the library. As Malpas suggests, "It's the responsibility of the library to have a long-range view of the relative importance of print in the research community and to be a few steps ahead of the faculty—it's the responsibility of the library community to lead" (Spiro 2009h).

In this period of transition, libraries may use pilot programs and interim solutions to ease the shift to electronic holdings. As *ITHAKA's 2006 Studies of Key Stakeholders in the Digital Transformation in Higher Education* recommends, libraries may want to pilot programs to shift to digital collections with more-receptive science faculty before launching general initiatives: "Moving aggressively to a digital platform in the sciences may not provoke much resistance,

but in the humanities may bring substantial faculty complaints" (Schonfeld and Housewright 2008, 17). In addition, several experts we interviewed suggested print on demand as an interim solution to the problem of reading long-form material. Print on demand can quickly and inexpensively deliver a digital file in a readable format. Through Springer MyCopy (Springer 2010), for instance, patrons at academic libraries that have purchased Springer e-book collections can easily order a softcover version of a Springer e-book for $24.95. Likewise, the Espresso Book Machine, developed under the leadership of publishing veteran Jason Epstein, takes a digital file and rapidly prints, binds, and trims a good-quality paperback. Currently the Espresso Book Machine is available at University of Michigan Library, McGill University Library, Bibliotheca Alexandrina, New Orleans Public Library, and the University of Alberta Bookstore, among other locations (OnDemandBooks 2009). Although print on demand may be an attractive interim solution, it does not provide the searchability and manipulability that characterize electronic text. Moreover, ink and paper are consumed to provide each reader with a copy of the book; the same copy is presumably not used by multiple readers, as is a library book.

Ultimately, what matters most to researchers is whether they are able to work productively. Bruce Miller reports that researchers at Merced have embraced the new model library because they get access to the resources and services they need. Indeed, they do not recognize that the library is any different from the traditional model: "They don't need to know how we do it, but whether they are getting what they need, and they are" (Spiro 2009i). Furthermore, the entrepreneurial faculty who seek appointments at new universities may be more open to new technologies and approaches. As Steve Shorb noted, the new library at NYUAD will be "different in nature, but then our institution is different. … By nature, the faculty will be interested in doing new and innovative things" (Spiro 2009f).

2.2.6 Resistance by librarians

Another obstacle to an all-digital library may be librarians themselves. As Bruce Miller suggests, the most significant challenge in shifting from an old to a new model library is social: convincing professionals who are good at what they do to do something new, and training them to adopt new roles. "How do you walk in and ask them to do something new? ... How do you make the transition?" (Spiro 2009i). Historically, libraries have been conservative institutions focused on preserving the past. Librarians may resist change because they are accustomed to doing things a certain way, worry about losing their jobs, or fear change, or because their organization is rigid and static (Weiner 2003). When Drexel, one of the first libraries to embrace e-journals, programmatically began replacing print with e-journals, resistance from some collection development librarians posed the "primary obstacle to implementation" (Montgomery and Sparks 2000, 16).

Librarians may have good reasons for thinking that it is

premature to shift to an all-digital library. As Sue Polanka observes, adopting e-books is "a major disruption to workflow" (Spiro 2009g). Whereas librarians can easily order print books, they may have to consult an attorney to review the licenses associated with each e-book or e-book collection that they order. As libraries increasingly collaborate in sharing their resources, librarians may reject the rights restrictions that come with e-books. Some may also worry that our common cultural legacy will decline if libraries replace print collections with electronic, given that how to provide long-term access to collections remains an open question. Insisting that libraries' most important job is providing access to and preserving books, Thomas Mann argues that the print book best communicates wisdom and promotes understanding of complete ideas as opposed to bits of information. He contends that libraries can maintain a strategic advantage in the electronic era by providing access to resources otherwise hard to come by, such as print books and licensed databases (Mann 2001).

Yet librarian resistance may be overstated and can be overcome. According to the 2006 ITHAKA survey, librarians, particularly those at larger institutions, view licensing electronic resources as an important function, suggesting that "leading-edge libraries are beginning to change their priorities to match those of faculty and students" (Schonfeld and Housewright 2008, 6). As Schonfeld comments, based on the 2006 ITHAKA survey, "The collection development directors in 2006 were ahead of faculty members in terms of their commitment to and enthusiasm for moving away from print" (Spiro 2009k). Librarians' resistance to change can be overcome through communicating clearly, involving stakeholders in decision making, providing training and development opportunities, and designing jobs appropriately (Farley, Broady-Preston, and Hayward 1998). For instance, Drexel surmounted librarians' reluctance to transitioning to electronic journals by collecting and evaluating relevant data, being flexible, and emphasizing leadership's commitment to e-journals. Data provided by surveys such as LibQUAL+ and the ITHAKA faculty surveys, initiatives to collect usage statistics such as COUNTER, and decision-making tools such as the Print Collections Decision-Support Tool may ease some of librarians' anxieties about migrating to electronic formats because it enables them to base decisions on solid evidence (Association of Research Libraries 2009; COUNTER; ITHAKA S+R). Miller avoided social resistance by personally hiring every librarian at UC Merced—something that the directors of other new academic libraries could likewise do.

2.2.7 Difficulty integrating e-books into library workflows

According to a recent survey of ARL member libraries, even the libraries that have moved farthest in adopting e-books view the environment as "evolving," given the diversity of publication and access models, uncertainty regarding preservation, and rights issues such as the inability of libraries to loan e-books (Anson and Connell 2009). Some libraries have been reluctant to adopt e-books because of

the heterogeneity of pricing and business models, DRM restrictions, the lack of a standard file format, and limitations in discovering and delivering e-books (Ball 2009; Connaway and Wicht 2007). The infrastructure and market for e-books are constantly changing (Ball 2009, 19). Although some standard e-book formats such as EPUB are beginning to emerge, there is still significant flux and divergence from those standards. Standards are important in enabling consumers to read content from multiple publishers on their devices, to move content around to multiple devices, and to preserve books for the long term.

The ever-shifting e-book market bewilders many librarians. A 2006 JISC study found a great deal of ignorance and confusion about different e-book acquisition models (Higher Education Consultancy Group 2006).[12] Likewise, a 2007 survey of librarians sponsored by ebrary revealed that "80% of the respondents found e-book acquisitions models confusing" (Connaway and Wicht 2007). Michael Kronenfeld characterizes the e-book market as a "mishmash" without clear business models, noting that some publishers require libraries to buy books in packages, others as individual titles, and still others as subscriptions (Spiro 2009m). In acquiring e-books, librarians must assess criteria such as DRM, the availability of MARC records, the number of concurrent users, archiving policies, and flexibility in assembling packages (Cleto 2008). Typically publishers do not want to sell single e-books; they prefer to offer batches of books through aggregators. Yet licensing terms that require librarians to buy e-books in packages limit their ability to customize collections according to the needs of their research community. Libraries—and ultimately end users—must deal with a multiplicity of platforms, which makes both ordering books and navigating user interfaces difficult. There is also confusion over whether to lease or buy an e-book collection. Leasing is typically less expensive and means that the library will not be stuck with the DRM regime adopted by the aggregator. However, purchasing helps address concerns about long-term access to a collection (Spiro 2009a). Publishers and librarians will need to work together to address these concerns and to streamline the process for ordering, managing, and making available e-book collections.

2.2.8 Concerns about preservation

For both librarians and researchers, a key concern about moving to digital collections is how to uphold the library's mission to preserve scholarly materials for the long term. Books printed on acid-free paper can last for hundreds of years, while the life span of digital information is, as Jeff Rothenberg jokes, "forever—or five years, whichever comes first" (Rothenberg 1998). As Robert Darnton suggests, "The best preservation system ever invented was the old-fashioned, pre-modern book" (Darnton 2008). Libraries have preserved print

[12] To help libraries clear up the confusion, JISC developed an Academic Database Assessment Tool (http://www.jisc-adat.com/adat/home.pl) that compares e-journals, databases, and e-book platforms according to features such as search, usage restrictions, indexing, access control, and metadata.

books for hundreds of years, but they fear that digital resources may not last a generation. What will guarantee long-term access? What happens if the data are lost? Without legacy collections, new libraries may not bear the same expectation to preserve large print collections, but they still share responsibility with established libraries for ensuring long-term access to scholarly information in a range of formats. Yet work is under way to develop ways to preserve digital information, including technical approaches such as migration and emulation, certification and auditing of digital repositories by organizations such as the Center for Research Libraries (CRL), and collaborative enterprises among publishers and libraries such as Portico and Lots of Copies Keep Stuff Safe (LOCKSS).

Digital preservation is both a technical and a social problem. Data formats quickly become obsolete, as do the devices on which digital files are played or displayed. Libraries are exploring technical solutions such as migration (moving bits from one format and device to another) and emulation (simulating the environment in which the original files were played). Along with the technical challenges of preserving digital information come the institutional problems of ensuring long-term access, particularly when works reside on publishers' servers rather than on library shelves. Libraries have long institutional histories, whereas a company such as Google or Elsevier could go out of business or change its business model at any time, taking its collections with it. Who will be responsible for preserving digital content 20 years from now? Who will pay for it? How will libraries have trust in the system for preservation? Libraries may expect publishers or other organizations to take on the responsibility for providing long-term access to electronic resources. In justifying Drexel's decision to replace print with electronic journals, Montgomery and Marion argue that "maintaining a print collection takes significant resources that can be better used elsewhere," and that "appropriate organizations would step forward and assume responsibility for archiving these journals" (Montgomery and Marion 2002, 97). By opting for subscriptions to e-journals and monographs over print, libraries cede a degree of control to the publisher. If libraries "rent," rather than own, scholarly resources, they worry that they will lose access to content if they can no longer afford the subscription cost or the publisher changes the terms of access or ceases publication. If the publisher fails, who will take custodianship of its content? Libraries should move cautiously, making sure that they are aware of the terms for postcancellation access. Libraries also need to guarantee that the content they deem important, including back files, is archived through initiatives such as LOCKSS, CLOCKSS (Controlled LOCKSS), and Portico, where users will still have access to back issues following the cancellation of a title.

As Yogi Berra said, "It's tough to make predictions, especially about the future," but libraries and publishers may be able to shape that future by taking responsible actions now. They have successfully faced past preservation challenges, such as extending the durability of print books by shifting to acid-free paper. At present,

there are no set solutions to digital preservation: much research remains to be done in this area. Solutions to digital preservation will likely emerge from collaborations among libraries, scholars, funding agencies, and publishers such as Portico, LOCKSS, DuraCloud, and MetaScholar (Spiro 2009e). Commercial organizations such as Oracle, IBM, and Amazon, which are beginning to offer cloud storage solutions for managing content, also have a stake in ensuring that these digital resources are preserved. Portico, a not-for-profit digital preservation service, works with publishers to ingest and archive their electronic content, normalizing the content to a "standardized archival format" and ensuring long-term access to the journal (ITHAKA 2005). By participating in Portico, libraries retain campus-wide access to archived journals that they subscribe to should a trigger event occur, such as a publisher ending its operations or its publication of a journal. Subscribers to Portico will have access to their past subscriptions provided they continue to subscribe to Portico.

While Portico adopts a centralized model for preservation, LOCKSS takes a distributed approach. LOCKSS provides the core technology and infrastructure for harvesting and replicating journal content across several institutions that have common subscriptions to a title. This distributed approach allows the integrity of each copy to be constantly checked and compared with that of other copies of the same title in the distributed network. LOCKSS will automatically repair a copy that has been corrupted based on the majority of the copies with an identical representation of the item (LOCKSS 2008a). Through LOCKSS, participating libraries have access to the back issues of journals they have subscribed to if they decide to cancel their subscription in the future. More than 300 academic publishers permit their content to be preserved by LOCKSS, and more than 200 libraries around the world belong to the LOCKSS alliance.

In addition to LOCKSS, CLOCKSS provides a "dark archive" where participating publishers' content is comprehensively archived across a limited number of geographically distributed networked library locations worldwide. The CLOCKSS archived content is kept inaccessible until the occurrence of a "trigger event," such as the publisher going out of business or a massive long-term technical failure on its servers (CLOCKSS 2008). If a trigger event occurs, access to the journal's content will be made available through a third-party provider and the materials will be accessible on the Web to all; subscription to CLOCKSS is not a requirement for access to the content when there has been a trigger event.

Although these initiatives first focused on journals, in 2008 Springer launched a pilot project to investigate the feasibility of preserving e-books in CLOCKSS (LOCKSS 2008b). Likewise, Portico is preserving e-books, including works published by Elsevier/ScienceDirect, Duke University Press, Walter de Gruyter, and Springer (Kirchhoff 2009; Portico 2009). Libraries are also collaborating with each other on preserving digital collections. For instance, the HathiTrust, a collaboration among more than 20 research libraries, aims to construct a cooperatively managed digital archive of

digitized works (including works scanned through the Google Books project); enhance access, particularly for participating institutions; preserve the collection; and develop shared storage strategies (Hathi Trust 2010).

Preservation strategies for digital content continue to emerge with the increased recognition that such content is vulnerable and must be sufficiently replicated to support reliability should a digital copy disappear or become corrupt. One such strategy is the MetaArchive Cooperative, where institutions are participating in a collective to archive each other's institutional repository content in a distributed manner using the LOCKSS technology (MetaArchive Services Group 2009). Another example is DuraCloud, a service offered by the DuraSpace organization that enables subscribers to have their institutional repository content replicated across the cloud storage environment, with DuraSpace serving as the broker for cloud storage among several commercial cloud storage providers (DuraSpace 2009). Additionally, the DuraCloud offering would allow institutions to share their content with other repositories, making it easy to access digital content at several institutions. This addresses concerns that have been expressed by users about the inability to easily exchange information between digital repositories, a restriction that inhibits research that relies on these digital assets. The Integrated Rule-Oriented Data System (iRODS) approaches preservation and exchange of information in digital repositories from a policy-based data management perspective. With iRODS, management policies can be enforced across systems as data are exchanged between repositories (Rajasekar et al. 2010). The Open Archives Initiative–Object Reuse and Exchange (OAI-ORE) protocol is starting to be implemented in repositories to facilitate the exchange of items, helping support long-term preservation of digital assets (Open Archives Initiative). The PREMIS (Preservation Metadata: Implementation Strategies) standard for preservation metadata continues to evolve, creating a framework for capturing critical preservation metadata associated with digital objects (Caplan 2009). Collaborative organizations such as HathiTrust and the Open Content Alliance are forming to provide stewardship and preservation for the digital resources resulting from mass-digitization activities like the Google Books scanning project. National libraries are also taking responsibility for digital preservation. For instance, in the Netherlands, Koninklijke Bibliotheek, the national library, has committed to preserving Dutch e-journals (Koninklijke Bibliotheek 2009). The European Union has funded multiple projects to develop digital preservation tools, standards, and communities. These include the PLANETS Research Project (Planets 2007), CASPAR (CASPAR Project 2009), and Digital Preservation Europe (DPE 2009). Although these initiatives demonstrate that the library community is beginning to tackle the problem of preservation, concerns remain about how to ensure long-term funding for digital archiving, who has responsibility for preservation, how to evaluate whether an archive can be trusted, what to archive, how to control versions, and how to deal with the technical

challenges of maintaining long-term access to data (Hunter 2007). However, the Center for Research Libraries' Certification and Assessment of Digital Repositories program is evaluating repositories such as Portico using the Trustworthy Repositories Audit and Certification Checklist (TRAC), thus assuring the community that certified digital repositories can be trusted (Center for Research Libraries). Tools such as the Digital Repository Audit Method Based On Risk Assessment (DRAMBORA) provide a means of performing a self-audit for institutions running their own repositories (Digital Curation Centre 2010b). It is important that libraries, as well as funders, recognize the investment needed to bring digital preservation approaches to maturity and incorporate them into the normal operations of the library so that access to digital resources will be assured.

In addition to developing strategies and technologies for digital preservation, libraries are developing cooperative methods for preserving print. As libraries move to digital collections, many are exploring withdrawing print copies to save space and money. At the same time, they recognize that at least some copies of the print must be preserved to address concerns about the quality and long-term availability of the digital version, as well as to negotiate campus politics. Furthermore, some researchers will need to consult the print version of a work to study its physical characteristics. ITHAKA's *What to Withdraw: Print Collections Management in the Wake of Digitization* takes a systematic look at the preservation of print scholarly journals, setting "time horizons" for print preservation based on the belief that "many of the rationales for retaining print are likely to decline over the course of time" (Schonfeld and Housewright 2009, 2). It lays out criteria to consider in determining whether copies of the print journal need to retained, including whether the journal is text-only or has images, the quality of the digitization, how well the digitized version is being preserved, the "reliability" of the "digital provider," and "campus politics" (Schonfeld and Housewright 2009, 2). In the case of "well-digitized digitally preserved text-only materials," the report recommends that at least one print copy be preserved if it is to be available in 20 years (Schonfeld and Housewright 2009, 2). To avoid risk and provide access beyond 20 years, at least two "page-verified print repository copies" should be preserved (Schonfeld and Housewright 2009, 2). Libraries can collaborate to preserve copies of printed books and journals in shared regional repositories such as Five Colleges, Inc. (Massachusetts), the Committee on Institutional Cooperation (CIC) libraries, and PASCAL. Through distributed shared print-preservation networks, libraries can evaluate what resources are already being preserved and decide whether local copies can be discarded (Payne 2007; Reilly 2003). The consensus seems to be emerging that preservation should be a collective responsibility, but it remains to be determined how participants in such an effort can best work together and distribute responsibilities and costs. Libraries need to be convinced that investing in a common preservation infrastructure benefits everyone (Spiro 2009h).

2.3 Economic Considerations

The role of libraries as intermediaries in providing a public good for the communities they serve has been a key argument in defending the high budgets for these organizations, especially for those in higher education. For research universities, the quality of scholarship has historically been tied to the quality of the research collections maintained by the campus research library (Courant 2008). Libraries provide ready access to the scholarly record not only by carrying the significant works that support the scholarship of the university but also by offering the services that enable the discovery of relevant materials within the millions of volumes housed in the library stacks. With the shift from print to digital in both producing and consuming scholarship, the value of libraries will be increasingly scrutinized to understand their contribution to the advancement of the university's mission. At research universities, libraries have multimillion dollar annual budgets. With economic downturns such as those experienced in 2008–2009, this large number stands out when universities are looking to trim their budgets. It will be critical for libraries to continue to demonstrate their relevance beyond simply subscribing to scholarly materials. Libraries have increasingly become collaboration and study spaces as users more frequently rely on digital resources than on print. The cyberinfrastructure of services, hardware, software, and training necessary for supporting scholarship must be available to make full use of the digital content. The virtual environment is quickly replacing the bricks-and-mortar library. This presents an opportunity for increased collaboration among libraries, since specialized, distributed services can be provided more effectively and with greater cost efficiency than replicating these services in each physical location. This section examines issues that are affecting the economic feasibility of achieving a primarily digital library by looking at statistical trends in research libraries, shifts in staffing with digital resources, changing roles and responsibilities, new services with digital resources, changing business models, open access, and the requisite cyberinfrastructure.

2.3.1 Resource expenditures

The ARL releases statistics each year that show the expenditures of its 123 member libraries, broken down and aggregated to give a comprehensive picture of where the funding is going. The most recently published statistics reflect expenditures through the 2007–2008 academic year (Kyrillidou and Bland 2009). For 2008, the median overall budget of ARL libraries was $24.8 million, with 12.59 percent going toward operating costs, 45.43 percent toward staff, 0.61 percent toward contract binding, and 41.37 percent toward library materials. The median materials budget during this time was $10.5 million. Electronic materials expenditures were 53 percent of the total materials, a median value of $5.4 million. The total materials budget for 2008 increased an average of 5.8 percent over 2007; electronic resources expenditures rose 18.16 percent during this same period. Since the 1993–1994 academic year, electronic resource expenditures

have shown a significantly greater percentage increase compared with overall materials budget increases, as shown in figure 1 (Kryillidou and Bland 2009, 19).

During the 15 years of data reported, the average annual percentage increase for electronic resources was 26.47 percent, while the average annual percentage increase for all materials was 5.97 percent. Electronic resources include not only electronic serials but also computer files, bibliographic utilities, hardware and software, and document delivery/interlibrary loan (Kyrillidou and Bland 2009, 18). Since 2000, 88 to 90 percent of the electronic materials expenditures have been for serials, with serials expenditures overall far outpacing those for monographs. Figure 2 compares increases in serials expenditures with increases in monograph expenditures since 1986 (Kyrillidou and Bland 2009, 11). The trend toward increased e-resources expenditures is clearly indicated in the statistics.

Staffing considerations also affect the economics of working with primarily digital resources. In 2008, ARL libraries employed, on average, 260 staff. Approximately 40 to 50 percent of these staff were support staff rather than professionals (Kyrillidou and Bland 2009, 62). As libraries shift increasingly to electronic holdings, the number of support staff needed will likely decrease since there will be less need for manual processing of materials. The composition of professional staff will also experience a shift from experts being locally resident at each library to more distributed services, with experts available to serve a broader community of users. Selection and discovery services

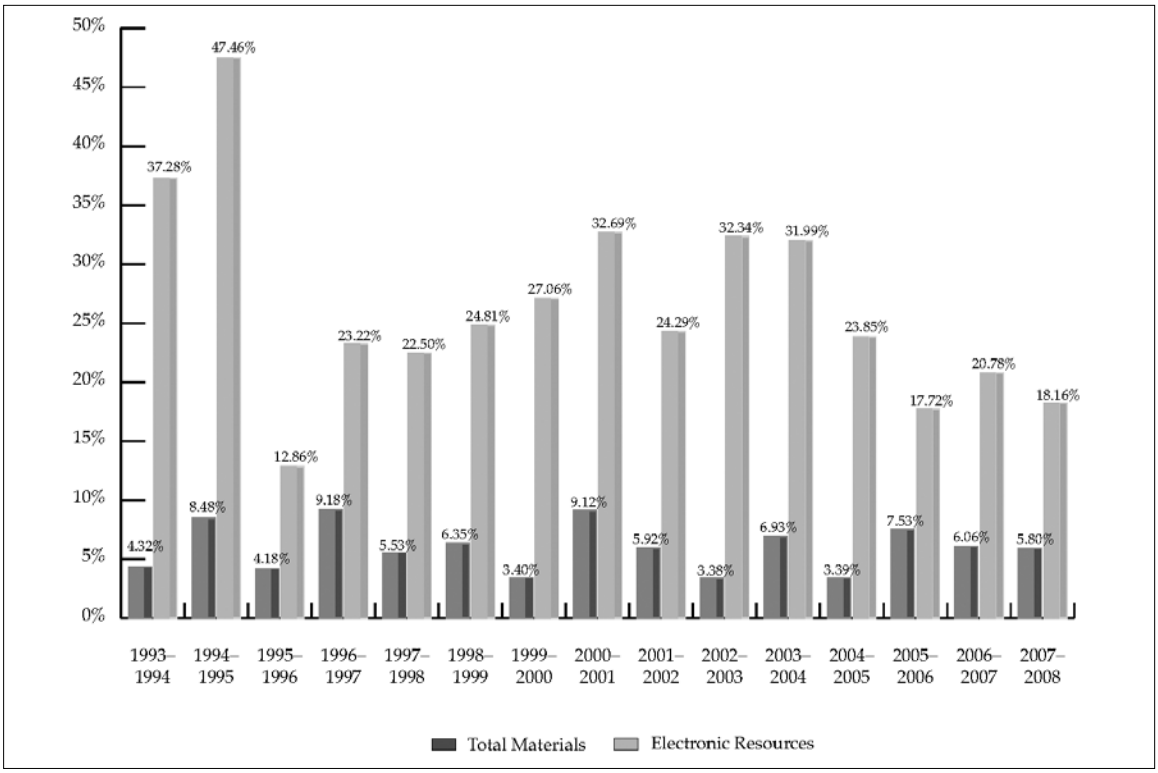

Fig. 1: Electronic resources versus total materials expenditures, 1993–2008, yearly increases in average expenditures, ARL libraries. *Source: Kyrillidou and Bland 2009, 19.*

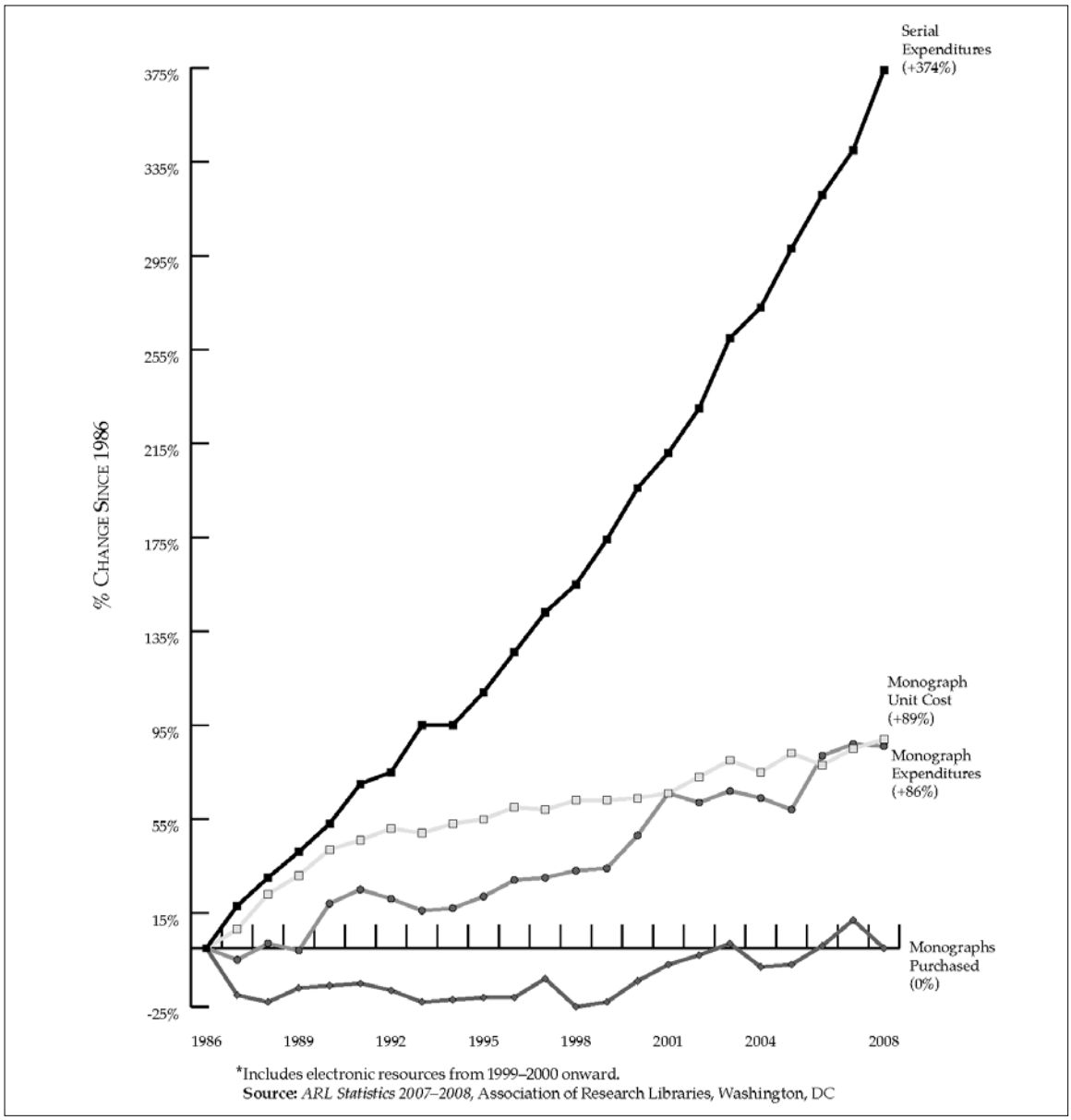

Fig. 2: Median monograph and serials expenditures at ARL member libraries, 1986–2008. *Source: Kyrillidou and Bland 2009, 11. Note: figure includes electronic resources from 1999–2000 onward.*

will still be required, but these services will include new offerings; working with e-resources changes the nature of the way new information is revealed when the knowledge exists in electronic format.

A 2006 study by Ellen Safley raises questions about the need for local selectors when working with e-resources. A comparison of the use of e-book collections prepared by librarians from the University of Texas system (NetLibrary) with that of a collection prepared by vendors (ebrary), revealed that the percentage of titles used in each collection was about the same. "Because of the similarities between the subject usage of NetLibrary and ebrary, the importance of librarian selection rather than subscription purchase is questionable and further testing is needed," the author concluded (Safley 2006, 453).

Vendors can hire librarians to create the collections they offer to libraries. These selectors are important for ensuring that the quality of the prepared collections offered to libraries is high. As Safley notes, more analysis is needed to better understand the need for local library staff versus vendor selectors; staff reductions at the local level would lead to savings, but this economic benefit must also ensure that the quality of the collection remains high.

Others have experimented with patron-driven selection of titles over collection building by professional librarians. A report on a program at Purdue University in 2002 indicated that 68 percent of patron-driven acquisitions circulated at least once after their initial use compared with 36 percent of titles normally acquired by the library (Anderson et al. 2002). As libraries increasingly outsource selection of materials, they will play a greater role in curating digital content both from their own institutions and from other sources to obtain the digital resources their patrons regard as critical. In 2007, David Lewis predicted that "in the next 20 years, less than 50 percent of a library's collection-related investments will go into purchasing collections and over 50 percent will go into curating digital content" (Lewis 2007, 427). This predicted shift of library investment from collection purchasing to digital curation remains to be seen, but the increased attention to maintaining, preserving, and improving digital assets suggests that digital curation is quickly becoming a core activity for libraries. Organizations such as the Digital Curation Centre in the United Kingdom are providing training and expert knowledge in this area, with tools such as DRAMBORA emerging to support libraries in their digital curation efforts (Digital Curation Centre 2010a).

2.3.2 Cataloging and metadata records

With print resources, the library catalog, along with the services of subject librarians, were the critical keys to finding information. With e-resources, traditional librarian roles are being challenged. Information discovery is now enabled by search engines and shared online catalogs such as WorldCat. Full-text searching is also possible, enabling the discovery of resources that have not been specifically cataloged with the terms that would otherwise allow them to be found. There has been much debate over the quality of vendor-supplied catalog and metadata records, but libraries are increasing their purchase of these as a cost-effective means of creating records for their holdings. It is doubtful that libraries will be able to sustain the workflows involved with creating the highest-quality catalog records for their resources.

> The time and energy required to do Library business is unsustainable. We have people performing duplicative work throughout our system. We are unable to share matching resources or records across our multiple catalogs, content management systems, and differing standards. These redundancies have opportunity costs in terms of services we do

> not have the time or staff to offer. We all agree that the cost of our Bibliographic Services enterprise is unsupportable as we move into an increasingly digital world, yet a solution is nowhere in sight (Declerck et al. 2005, 9).

For e-resources, the descriptive information (for example, MARC record or metadata) can often accompany the publication, with professional catalogers working for publishers to create high-quality metadata that can be shared. The University of California's Bibliographic Services Task Force recommended that UC catalogers accept vendor-supplied records as provided without adding enhancements. "We must adapt and recognize that 'good enough is good enough', we can no longer invest in 'perfect' bibliographic records for all materials" (Declerck et al. 2005, 25). Many academic libraries have already started to experience a decrease in technical services staffing with the availability of electronic records and fewer physical volumes to handle. The vendors of electronic resources also have a strong incentive to employ expert subject catalogers to ensure high-quality metadata for the published e-resources so that their customers are more accepting of their products.

Though librarians have expressed concerns about the quality of vendor-supplied bibliographic records, the enhancements that abstracting and indexing services can provide to the traditional catalog record can help users discover information. This also provides an opportunity to reconsider approaches to cataloging that would encourage partnerships, rather than competition, between libraries and electronic resources vendors. In response to recommendations by the Library of Congress Working Group on the Future of Bibliographic Control to eliminate redundancies and improve efficiency by using information from network resources, Deanna Marcum, associate librarian for library services, stated, "Cultivating partnerships for the exchange of data may offset the expense of purchasing metadata from aggregators, for example, exchanging controlled data such as authority records and controlled subject headings for descriptive metadata" (Marcum 2008, 10). As she suggests, newer models whereby professional catalogers share the responsibility of creating complete, very high-quality metadata for resources could lead to a more distributed approach with focused expertise. Ensuring the bibliographic integrity of scholarly works can be done by a collective, trusted organization and will require interinstitutional collaboration (Courant 2008, 27). Currently, OCLC does provide some of these types of services for a fee (OCLC 2010).

In addition to bibliographic records for discovery of information, e-resources enable greater computational approaches to analyzing content; this requires skills and services that support the use of technologies such as data mining and document clustering. As library expenditures continue to shift toward more digital content acquisitions, staffing transitions will follow with the services needed to work more effectively with these resources. Figure 3 shows trends in ARL member expenditures from 1986 through 2008 (Kyrillidou

and Bland 2009, 15). The greatest increase in expenditures by far is in library materials (286% over the 12-year span), with serial expenditures showing the greatest climb (374%). Given that most e-resources in current library collections are serials, this raises the question of whether or not monographs will shift to this same trend. During the 12-year period, monograph expenditures increased 86 percent, which was lower than the consumer price index (CPI) increase of 101 percent for the same period. Staff salaries have consistently exceeded the CPI during this time, but have shown an even greater increase since 1998. As the staffing profile shifts to support new services requiring more computer science and informatics expertise, salaries are likely to increase, but overall staffing will decline because less support will be needed for working with print materials.

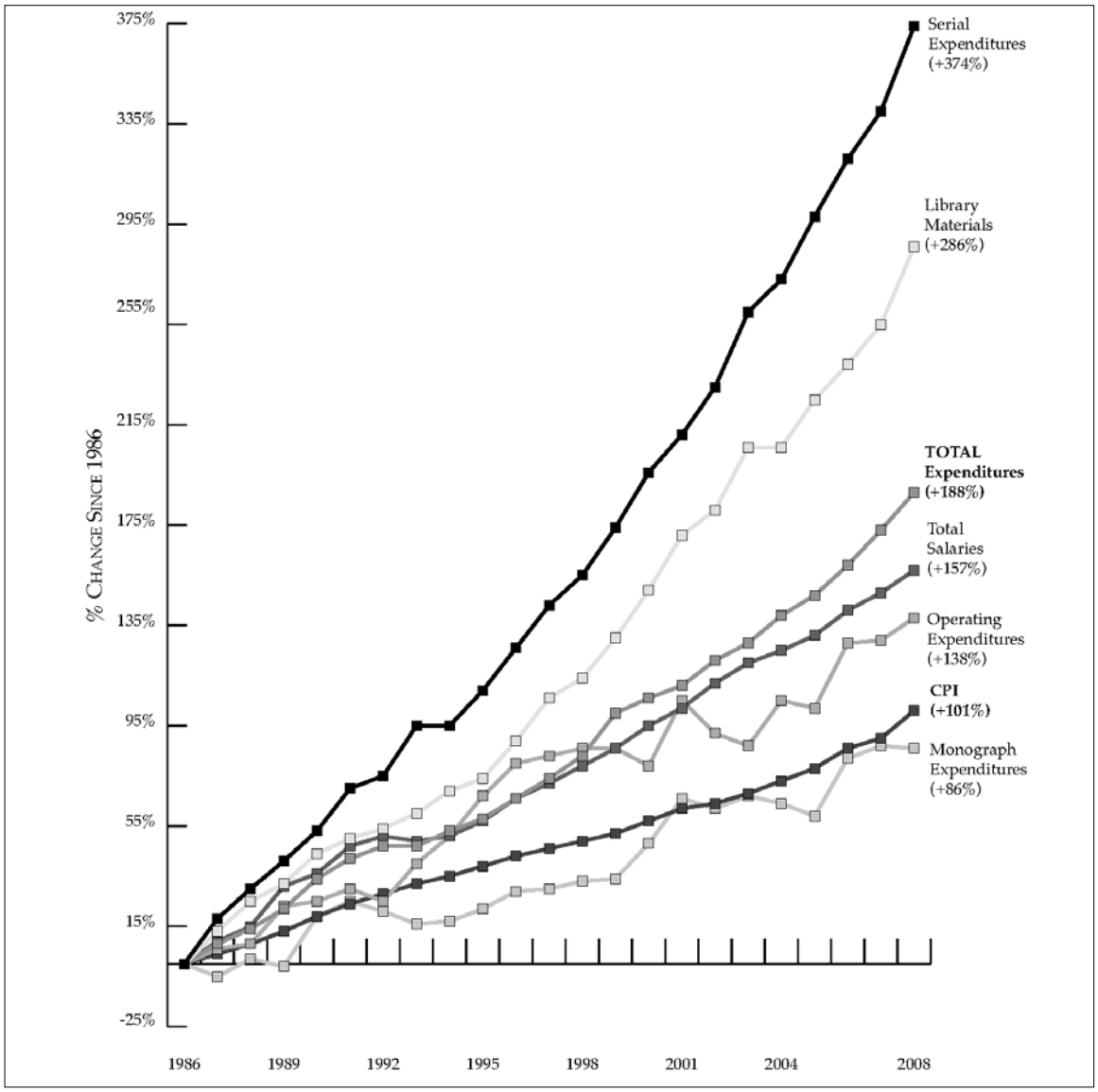

Fig. 3: Expenditure trends in ARL member libraries, 1986–2008.
Source: Kyrillidou and Bland 2009, 15.

2.3.3 Impact on quality of scholarship

It is difficult to measure the impact of the library's expenditures on the quality of scholarship produced by the institution. One approach might be to compare an institution's library expenditures with its research expenditures (e.g., research grant funding). One can argue that an institution's success in acquiring grants can be linked to the resources available to researchers in understanding their fields and staying abreast of new developments. There are other factors that influence the success of grant applications as well as additional correlations between the library's expenditures and the quality of the institution's scholarship, but research expenditures are offered as a tangible means of beginning to assess the impact of the library on scholarship. While the ARL annual library resource expenditures are readily available, there is not a comparable collection of annual research expenditures by these same institutions. It is possible to gather such information for each of the institutions, and we hope to do so in follow-on research. Some university libraries have calculated a return on investment (ROI) where they have considered the research expenditures at their institutions. "A reliable ROI would answer the question of how much quantifiable value the University received for every dollar it invested in the library" (Kaufman and Watstein 2008, 2).

By examining the trends in research expenditures and comparing them with the trends of ARL library expenditures as a whole, library materials expenditures, and library expenditures on e-resources over time, it is possible to draw some conclusions regarding a correlation between research expenditures and the investment in library resources. The University of Illinois has methodically examined this information and concluded the following for fiscal year 2006:

- Over 78% of tenure-system faculty holding grants in 2006 used citations to the scholarly literature in their proposals.
- Over 50% of grants awarded to our campus came from proposals that included citations to materials accessed through the library.
- The average grant income at Illinois is approximately $64,000.
- Using those numbers in our formula, we arrive at an average amount of grant income generated through the use of library resources of just over $25,000.
- Multiply this average amount of grant income by the number of grants awarded in 2006 at Illinois, and divide that number by the total library budget during that year, and you arrive at an ROI of **$4.38** for every dollar invested in the library. (Kaufman and Walter 2008)

As library budgets increase, libraries will increasingly come under pressure to demonstrate their value to the campus. The University of Illinois' work to demonstrate the ROI may serve as a model that other libraries can follow for calculating their associated value. One statistic may be of interest to those who attempt such an effort: in 2008, U.S. federal research expenditures totaled $54.7 billion, a decrease of 2.5 percent from 2007. Materials expenditures among ARL libraries increased by an average of 9 percent during that same

period. The ROI approach to demonstrating value in library expenditures is not without controversy. Lown and Davis argue that attempts to apply ROI models across libraries fail to consider local differences that affect this estimate. The University of Illinois ROI study is faulted for not considering social value or increased productivity that resulted from the use of library resources (Lown and Davis 2009).

Materials costs have continued to increase, but publishing costs have changed with the shift to digital publications. The public good notion of "produce once, make available for all" works much better in a digital environment than in print. The previous economics of publishing, whereby it was expensive to print and then distribute, no longer holds true with digital formats. Making a copy is "free," and print on demand alleviates the need for large print runs to justify the print setup. As Courant notes, it's cheap to produce a "pretty nice book" that is generally quite acceptable (Courant 2008, 25). The cost of distribution is free when publications are electronic. Publishers have, however, found a successful business model that allows them to continue to increase prices for serials, even with the reduced printing and distribution costs. Libraries subscribe to journals that publishers maintain on their own servers, thereby making libraries dependent on the publisher for continued access to the content. Publishers have argued that the costs associated with editorial services and maintenance of the technology have not allowed them to lower their price

The business models for monographs, however, are still being developed. Unlike the market for journals, the market for monographs is not restricted primarily to libraries (Baker and Evans 2009, 167). Publishers have experimented with a variety of approaches to control pricing, including delaying the release of e-books for four months after the release of the print book (Rich 2009b). Approaches that have been used to sell e-books include using various subscription models, selling titles directly to libraries, selling access to collections, providing unlimited simultaneous access, allowing limited simultaneous access, restricting total usage of a title per year, selling collections by aggregators, providing limited-time subscriptions (much like interlibrary loan), and offering parts of books by "slicing and dicing" to provide only the relevant sections (Baker and Evans 2009, 169–170). A study at the University of Westminster compared the economics of a business model where simultaneous access was limited with that of a model that capped maximum annual access. The results showed that either business model would provide adequate access for the users (Grigson 2009, 8), although not limiting simultaneous access provided a better service model since the library could meet peak demands without exceeding the cap. Librarians responding to the HighWire press e-book survey indicated that the most acceptable business model for e-books is purchase with perpetual access, though other models are also acceptable (Newman 2010). Until the business models for e-books stabilize, it will be difficult to reliably plan expenditures for all-digital resources in the academic

library. Furthermore, it remains unclear, based on social, cultural, and policy issues, whether or not the shift to e-books will occur as soon as we would predict on the basis of business models alone. If users, especially scholars, find greater benefit in working in a hybrid environment, the shift could take longer and the business models could look very different from those we envision today.

The open-access movement has offered an alternative to high-priced journals controlled by publishers. Since open-access journals do not require libraries to formally subscribe to them, these journals have worked their way into collections as scholars have identified them as key resources for their fields of study. Open-access journals that have strong editorial boards and rigorous peer review are valued by their communities. As Conley and Wooders have noted, "The best advertising is through word of mouth by reputable scholars in an area. This comes automatically if the editorial board is a good one and can be helped with such things as self-funded conferences; Google Scholar and RePEc are also very good sources for advertising for open-access content. Remember, open-access journals are not trying to sell subscriptions, only to persuade people to submit papers and read the Content" (Conley and Wooders 2009). The value-added of editorial contributions by publishers is not as valued by authors in the digital environment as in the analog realm since authors in the former are able to control much of the presentation themselves. While authors would still like someone to "proof" their articles and readers would like editorial quality control, both are willing to compromise if it means significant cost savings at no sacrifice to the accuracy and reliability of content (Conley and Wooders 2009). Mandates by federal funding agencies requiring that publications resulting from federally funded research be deposited in open-access repositories have led the way to more freely available research information. The U.S. National Institutes of Health passed a mandate effective in April 2008 requiring deposit of publications into PubMed Central, where they are available as open-access documents (U.S. Department of Health & Human Services 2009). The Wellcome Trust in the United Kingdom passed a similar mandate much earlier requiring authors to place their publications in the repository now known as the UK PubMed Central (Wellcome Trust 2008). That mandate became effective in October 2006. Several colleges and universities have passed mandates requiring faculty to submit their publications to open-access institutional repositories (ePrints). Science and engineering communities have been quicker to embrace open access, with several starting their own open-access journals to protest the high cost of the traditional journals in their fields. Members of the physics community routinely deposit their prepublications as well as their final publications in the arXiv.org repository, started in 1991 (Cornell University Library). The repository has been so successful that the computer science and mathematics communities have also adopted it for their publications, with many of the papers citing publications from arXiv.org before they appear in the traditional journals that have accepted the papers for publication.

The long-term sustainability of open-access journals and repositories is not yet well-known since there are no long-term data or proven business models to indicate whether or not they can continue to operate into the future. The cost savings to libraries and colleges could, however, be significant if academic publishing continues to move toward open access. A distributed network of open-access repositories and journals maintained by colleges and universities around the globe would provide a means of sharing the costs for supporting open access to faculty research publications at participating institutions. Cornell University has requested that the top 200 institutional users of the arXiv.org service discussed earlier assist them financially with the continued costs of providing this service (Glazer 2010). It is unclear whether or not this same model can scale to other disciplines, but it is a useful approach in understanding the value, as well as the costs, of open-access scholarship. If these open-access publications can be fully accepted in promotion-and-tenure considerations, the journal publication model could shift away from the for-profit publishers in favor of the community-driven publications. In looking at the constantly rising expenditures for serials in ARL libraries, with a median expenditure of $7,097,140 in 2008 (Kyrillidou and Bland 2009, 10), this would result in substantial savings, even if the institutions provided funds to support the journals and repositories, since costs would be shared. We are a long way from realizing this utopian model of open access, to be sure. To make this shift would require research institutions to cooperate in supporting open access at a large scale, making academic publications genuinely a public good: produce once, and make it available for all (Courant 2008, 25).

2.3.4 Cyberinfrastructure

The cost for the cyberinfrastructure needed to support an all-digital library is a significant consideration in moving away from print. While the term *cyberinfrastructure* has been used to refer to various elements in a technology environment, we are defining it in this context as the combination of technology (hardware, software, and utilities), services, research, and training needed to support an all-digital library. New forms of information, such as multimedia scholarship and data sets, will place increasing demands on technology. Resources required to support the fully digital environment include more computers and storage to support digital content, especially large scientific data sets, with sufficient processing available to meet user demands. Networks must be robust and ubiquitous to enable users to readily access and interact with stored information. High-availability configurations to ensure reliability and sustainability of the server, storage, and network technologies will require redundancy in systems so that the information is available when there are technology failures. E-readers and other viewing hardware must be available to users; during the transition to digital, libraries may need to have these devices readily available to loan to users. Libraries may also be called on to provide equipment for creating the newer forms

of scholarship beyond text. Multimedia equipment, including audio, video, and camera gear, as well as specialized hardware needed for viewing or working with multimedia, will require an increased investment. All these technology resources will require increased power availability.

Software for managing large data sets, tools for supporting visualization, data and text mining applications, multimedia creation, software for music analysis, tools for supporting data preservation—these are just a few examples of the software applications that will need to be supported to effectively use and create the digital information for use in teaching and research. In 2007, the National Science Foundation's (NSF) Cyberinfrastructure Council identified four priority areas for supporting cyberinfrastructure: high-performance computing; data, data analysis, and visualization; virtual organizations for distributed communities; and learning and workforce development (Cyberinfrastructure Council 2007). While not all of these areas affect the library, most of them will in some way affect library staffing, technology, and support for realizing the NSF's vision. This needed infrastructure will require a greater investment over current library technology infrastructures, but there will also be an offset to costs associated with supporting millions of volumes of print titles. Many academic libraries currently rely on campus information technology (IT) organizations to provide core technology support, though some institutions have combined the library and IT operations. Strong partnerships between supporting IT organizations and libraries will be critical for supporting the shift to digital resources, and the budgets of these organizations will become more intertwined because of their critical interdependencies.

The additional cyberinfrastructure elements—services, research, and training—will affect libraries' staffing profiles. Shifting to digital content enables new means of exploring information to discover new knowledge. "What would you do with one million books?" This question launched the Digging into Data Challenge in 2009, with grants awarded to support a diverse set of projects working with massive amounts of digital content to discover new knowledge within these materials (National Endowment for the Humanities 2009). Music collections, speech data sets, image data, court proceedings, geospatial data, and manuscripts are examples of the data that will be explored. Automatic extraction of terms is possible, as are statistical analyses of the content that can help identify relevant information for researchers. Databases can be mined to discover patterns of interest. Visualization tools, along with the ability to tune algorithms to meet specific researcher needs, are among the services that libraries will be expected to provide as they acquire more e-resources. Staff who are knowledgeable about the tools available for analysis and who have deep subject matter expertise will be in high demand. As data sets are acquired, staff with backgrounds in science will be needed for data curation.

Libraries will need to hire more computer scientists and informatics professionals to work alongside subject matter experts and

information scientists. There are many research opportunities that must be addressed by the library community with massive digital content; among them are improved search algorithms, automatic classification approaches, document-clustering techniques, pattern recognition, and, perhaps most important, digital preservation. Working effectively in the all-digital environment requires comfort with and knowledge of advanced technologies. The library must also be prepared to provide training in these technologies so that users can work effectively with digital content. With the shift in skills and services, the shape of the organization will likely change, with much more reliance on virtual organizations to support the data cyberinfrastructure. Borgman et al. (2009) have discussed their effort to create a "cyberinfrastructure virtual observatory" that will enable effective data-management policies to be followed by virtual research organizations. As they point out, understanding data-based collaborations will have an impact on the design and development of digital library services and architectures.

Salaries of library staff will likely rise, given the required new skills. The number of staff, however, should diminish. Moreover, all the individuals with the needed new skills will not have to be stationed in a single library. Some academic libraries have already started to explore shared services, especially with common print collections. In October 2009, Columbia University Library and Cornell University Library announced plans to partner in providing shared services. The collaboration, named 2CUL for the acronyms of the two libraries, will initially focus on how to jointly transform their operations in three areas: "managing electronic resources and other nuts-and-bolts library work, building global-collecting capabilities, and creating a digital-preservation infrastructure" (Howard 2009c). New academic libraries that are primarily digital can implement these staffing and services requirements from the start, realizing an overall savings in staff costs and leveraging partnerships with other academic libraries. The shift is much more challenging for established libraries moving from print to primarily digital. The model of Columbia and Cornell to offer common services during the transition is helping these institutions bring in the needed expertise during the move toward more digital information.

Estimating the economic feasibility of having a primarily digital library at this time is difficult because of the many unsettled business and process models and the rapid changes in technologies. These factors are still shifting, putting budget planning at risk. However, as shown in figure 4, the trend toward increasing expenditures on electronic resources is undeniable. The percentage of the materials expenditures for electronic resources continues to increase, with a median of 53.06 percent of materials expenditures going toward these resources in 2008 (Kyrillidou and Bland 2009, 21). In a 2003 study, Connaway and Lawrence worked with 11 ARL librarians to estimate the resources needed for a paper library and those needed for a digital library (Connaway and Lawrence 2003). They considered four categories of resources (labor, space, materials, and equipment)

needed for all of the functions in the library and estimated significant savings in each of the categories. This study, however, was very limited in scope; more-extensive analyses are needed to demonstrate the true costs associated with digital versus print libraries. While it may be risky to planning for an all-digital academic library with many factors still changing, there is a much greater risk in not planning for this shift and continuing with business models and processes that have dominated the primarily print library.

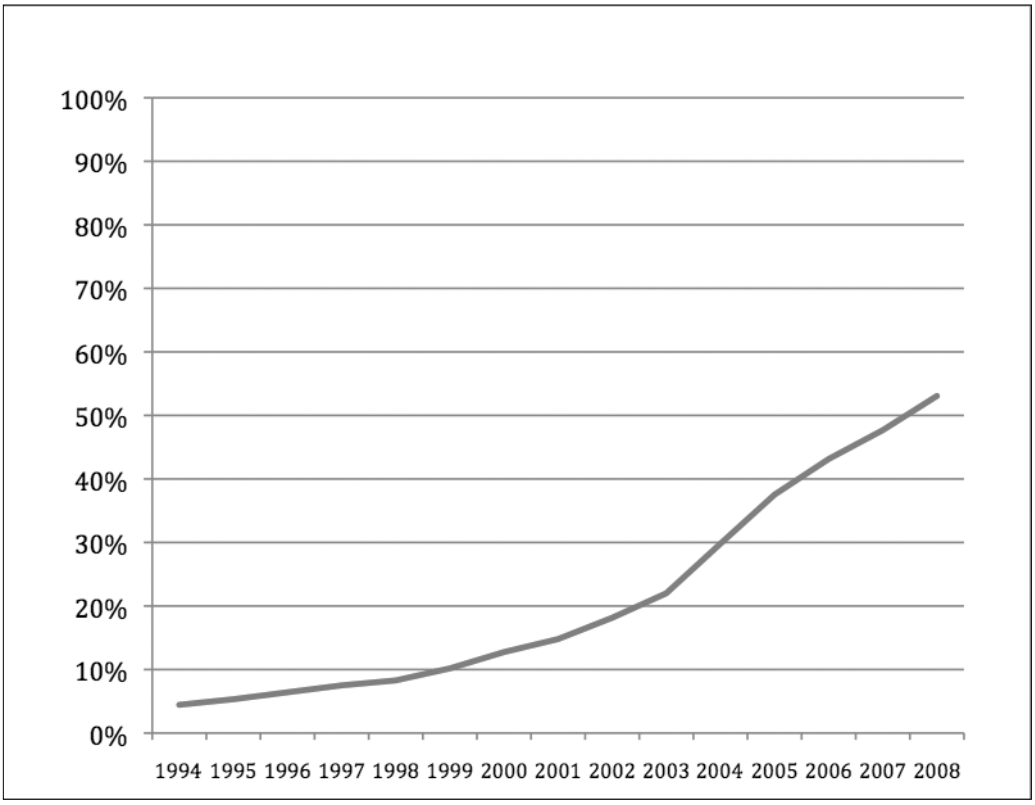

Fig. 4: Median percentage expended by ARL libraries for electronic resources as a percentage of total materials expenditures, 1994–2008.
Source: Kyrillidou and Bland 2009, 21.

2.4 Can a Research Library Offer All-Digital Collections Today?

Would it be possible for a new academic library to be primarily digital? Whether a new university library could rely on digital collections today depends on the type of institution, the disciplines it supports, its cooperative relationships, and its resources. In the case of a virtual university, its library could likewise be digital, since it would be inconvenient for students who were otherwise educated online to visit a physical facility to do library research. As Michael Furlough suggests, the ability to deliver information electronically will "become critical as we more fully accept online learning. If we do more learning on the Web, we have to deliver all content there" (Spiro 2009d). For-profit distance-education institutions such as the

University of Phoenix have worked out systems for the electronic delivery of most materials; what cannot be delivered online is mailed to the student's home (Spiro 2009h). For community and teaching colleges, a library could subscribe mostly to electronic journals, retaining perhaps a small print collection of core journals and books. If the institution were focused just on law, medicine, or some areas of science and engineering, it could offer primarily digital collections, but would probably need to have a small print collection of essential textbooks, design books, and core material that has not yet been digitized. Another variable is the extent to which the library is enmeshed in networks of collaborative relationships that enable it to borrow what it does not hold. If a library is part of a strong consortium such as the California Digital Library (CDL) or OhioLink, it would have easier access to distributed collections and might not need as many print resources itself.

For the purposes of this report, however, we are assuming that we are creating a library for a new research university in the United States. Most major disciplines are represented at this hypothetical university, although it does not have a law school, business school, or medical school. At this time, we believe that a new research library could provide primarily electronic access to journals and reference works, but it would need a core collection of monographs targeted to the teaching and research interests of the university, as well as a strong interlibrary borrowing program. In the humanities and social sciences, a critical mass of electronic monographs in these fields has not yet emerged, nor has a viable model for reading electronic versions of long-form works. There are, however, alternatives to an individual library building even a core collection. At present, many libraries are trying to comprehend the significance of networked information for their mission and experimenting with new models for providing resources and services. In 10 years or so, the environment may have changed so that libraries could depend primarily on digital collections, with the exception of archives, special collections, and a few print resources, such as artists' books, that may have difficulty making the transition to digital formats.

Our reluctance to declare without reservation that a research library could today go all-digital is motivated in part by a cautionary example. In 1995, California State University Chancellor Barry Munitz told *Newsweek* that the new California State University-Monterey Bay (CSUMB) would not have a physical library, saying, "You simply don't have to build a traditional library these days" (Hafner 1995). Dubbed the "21st Campus [of the CSU system] for the 21st Century," CSUMB opened in 1995 on a former Army base with a physical library in place. Of course, Munitz was premature in declaring that the time was ripe for a virtual library. The idea of not having a physical library was discussed and quickly dismissed after administrators determined that most materials, particularly books, were available only in print, and that the library building played an important role in the intellectual and community life of the university. Nevertheless, the CSUMB library did prefer electronic over print

delivery of information and focused on delivering "just-in-time" rather than "just-in-case" collections (Ober 2000). As John Ober, formerly the development librarian for electronic resources at CSUMB, observed, "While many libraries feel that electronic resources are an ever more valuable supplement to print, CSUMB feels that print resources supplement electronic access, and that they probably will become a less important supplement as time goes by" (Ober 2000, 126). Usage data from the 1990s showed that CSUMB students used electronic resources more than students at other campuses: while 72 percent of student library usage at CSUMB focused on online sessions and 19 percent on book circulation, the numbers were almost reversed at other CSU campuses across the state, where 71 percent of usage was for book circulation and 26 percent was for online sessions. Yet the library also embraced the importance of a physical presence. In December 2008, CSUMB celebrated the opening of its new, 136,000-square-foot library building. The facility features "skybox study rooms," an information literacy center, more than 65,000 books, a writing center, auditoriums, and a café. CSUMB exemplifies the idea of library as place, providing social spaces where people can learn, collaborate, and get help with their research.

As the CSUMB Library found, its "startup model must be modified to incorporate more print material, as access to electronic books is highly limited" (CSUMB Library 2009). The CSUMB Library argues that a significant print collection aids in recruitment of students, complements the idea of library as place, and may be necessary for graduate programs (CSUMB Library 2009). Even though CSUMB began with the bold vision of being a digital library and was ahead of many libraries in providing access to electronic resources, Library Director Bill Robnett suggests that it now resembles the other libraries in the CSU system, noting that "the more we changed, the more the older CSU libraries changed and moved toward us on the continuum" (Spiro 2009n). Thanks to the central CSU office negotiating deals for schools in the system, all the campuses have access to many of the same digital resources. Much has changed since 1995: electronic journals have reached the tipping point, students and faculty have become accustomed to accessing information online, and libraries have incorporated electronic resources into their workflows and begun experimenting with cooperative collection development programs. Indeed, Robnett believes that if an administrator today declared that a new academic library would be all-digital, it actually could be, assuming that the library had a sufficient budget and that it was located at a teaching-focused, rather than a research, university (Spiro 2009n).

3. Preliminary Case Studies of Twenty-First-Century Academic Libraries

As libraries confront the challenges of becoming twenty-first-century institutions adept at managing digital information, they can learn from the examples provided by academic libraries founded since

2000. New academic libraries can from the start prefer digital to print collections, design flexible spaces, and hire staff with strong technical skills and an openness to change. Although our hypothetical new library is at a research university, we can gain insights about twenty-first-century libraries by looking at a range of examples, including special libraries such as health science center libraries, libraries at small private colleges, and research libraries. To understand the strategies adopted by new academic libraries, we performed preliminary case studies of six academic libraries founded in the United States since 2000, including one public research library, two health sciences libraries, two libraries at private universities, and one library at a state university. In addition, we looked at one branch library for a campus of a U.S. private university that just opened in the Middle East. We conducted semistructured phone interviews with leaders of these libraries and gleaned information from the libraries' Web sites and publications.

To our knowledge, no previous study has systematically examined new academic libraries. We should emphasize that this research is preliminary and is based primarily on 30- to 60-minute interviews with library leaders. Long-term assessments of new academic libraries should be carried out to evaluate the challenges they have faced and strategies they have developed. Furthermore, our scope in this study is limited to just a few institutions in the United States (with the exception of NYU Abu Dhabi) and should be expanded to include universities across the world. However, even this preliminary analysis suggests that new academic libraries encounter common challenges and can offer valuable insights about the future of libraries. Donald Barclay, deputy university librarian at UC Merced, articulates some of the key questions that new academic libraries face—questions that likewise have relevance for established libraries:

> Should a twenty-first-century academic research library be organized along the lines of public services, technical services, special collections, and all the other traditional library divisions, or should its organization take some new form? Should the new research library have books or truly be an online operation? What form should the library building itself take? What kind and how many staff would be needed? How would staff provide reference services? If there were to be a physical collection, who would build it, and how? (Barclay 2007, 104)

In setting up a new academic library, a university has the opportunity to define its central mission, services, collections, physical space, staff, and online presence. At UC Merced, for example, the library focused on asking the right questions—not "Where are we going to put the reference desk?" but "How are we going to provide reference services?" As a result, Merced has no reference desk; it provides reference service through phone, e-mail, chat, and workshops. In the words of Library Director Bruce Miller, "We just get to start with what seems logical" (quoted by Carlson 2005).

Steve Shorb, library director for NYUAD, observes, "One of the

wonderful things about going to the Middle East and starting a new library is that you don't have to take all of the baggage. You want quality, but you get to start over" (Spiro 2009f). Whereas established libraries would need to convince library committees and faculty to transition toward the electronic, new libraries typically have more freedom in determining how they will deliver information and in shaping the expectations of users. New academic libraries could start with digital collections, at least for most reference works and periodicals. As Roger Schonfeld speculates, "You can see a lot of cultural discomfort among faculty members about making the shift from print to electronic that doesn't necessarily reflect their practices and habits. … If you don't have to make that shift but can simply introduce a new state, those concerns may be diminished" (Spiro 2009k). Furthermore, faculty who help launch new universities are likely more entrepreneurial and more open to change than faculty at existing universities. Likewise, new library directors can hire innovative, skilled staff who can focus on the future and are not limited by a fidelity to how things have always been done.

Even as new academic libraries have the opportunity to define themselves, they also face significant challenges. Most of the libraries we examined have small staffs, but they often serve a small student body. Whereas long-established libraries hold rich collections, including rare books and manuscripts, that have been built up over a number of years, new libraries lack such collections. As Roger Schonfeld says, "Libraries have tremendous assets, and some of the oldest libraries have some of the richest assets, such as special collections. It would be hugely expensive to buy these collections today" (Spiro 2009k). Established libraries are guided by traditions that give them a sense of identity and purpose, but many are also innovative in devising support for research and teaching, exploring preservation solutions, and collaborating with other libraries on building and maintaining collections. Moreover, they have figured out workflows that enable them to be productive, although these workflows might also be restrictive as libraries shift toward digital collections. While staff at established universities have long-established relationships with faculty, staff at new universities must create these relationships. However, many new academic libraries view these challenges as opportunities to focus on key priorities and develop innovative ways to deliver service. Below we profile seven new or emerging libraries: UC Merced; CSU–Channel Islands; Olin College of Engineering; Ikeda Library at Soka University; Arizona Health Sciences Library–Phoenix; A. T. Still; and NYUAD.

3.1 UC Merced

The UC Merced Library consciously departs from the traditional library model, taking as its motto "Not what other research libraries are, what they will be" (UC Merced Library 2009b). UC Merced, the tenth and newest campus in the University of California system, is located near a town of about 63,000 close to Yosemite. In 2005

Merced welcomed its first undergraduate class. As of fall 2009, Merced had approximately 3,190 undergraduate and 224 graduate students, but it ultimately expects to have 25,000 students (UC Merced 2009). According to Library Director Bruce Miller, UC Merced is the only research library to have been established in the United States since the 1960s, when UC San Diego, UC Irvine, and UC Santa Cruz were founded. Rather than serving as a warehouse for physical collections, the UC Merced Library aims to return to the "basics"—to provide access to information however and whenever faculty and students want it (Carlson 2005). In part, this approach is driven by the sheer economic fact that UC Merced could neither create a physical collection on par with that of most research libraries nor afford to hire a large staff. As Miller says, "The budgetary constraints kept us from imitating the traditional libraries." Currently UC Merced has 8 professional librarians, 9 staff members, and 39 student assistants. Yet Miller and colleagues also embraced the opportunity to create the twenty-first-century research library and challenge traditional library assumptions. UC Merced's physical collections are small—about 85,000 print books, as compared to 616,000 digital books—but as a member of the UC system it has rapid access to the more than 34 million items held by all of the member libraries (UC Merced Library 2009a). Indeed, Miller says that this access to the UC system's extensive resources enables the UC Merced Library to qualify as a research library, noting that "a book on the shelf at Berkeley or UCLA or San Diego is one of our books"(Spiro 2009i). A researcher who needs a book not at UC Merced can click a button on a Web page and have it delivered quickly, often by the next day.

In rethinking the traditional model of the library, UC Merced embraces several key principles:

- **Adopt a "container-neutral" policy**, providing the format appropriate for the use. UC Merced implemented a flexible policy for getting users access to collections: "If an electronic container is the best way to provide a particular information resource, then we provide it electronically. If a print-format container is best, we provide it that way. And if appropriate, we provide the same piece of information in both electronic and print-format containers" (Barclay 2007, 110).
- **Make collections patron focused**. According to Miller, UC Merced's collection budget is "upside down of conventional research university policies." Instead of divvying up the budget among each of the subject bibliographers, which limits flexibility, Merced first made allocations for electronic resources, deciding to buy all resources that UC colleagues deemed "tier 1" without further evaluation. Merced then set aside money to respond to patron demand, getting students and faculty almost anything they requested, whether by purchasing it, borrowing it through interlibrary loan, or paying for digital rights. Occasionally Merced will reject a request after consulting with the requestor, as when a faculty member asked for a $35,000 database, was shocked to discover the price tag, and canceled the request. According to Miller, faculty

members become "partners" in selecting content; they respond well when requests cannot be met because they typically do get what they ask for. Any leftover money is used to purchase contemporary monographs relevant to Merced's research and teaching areas. There is not a subject bibliographer on staff; rather, collection development is done "in aggregate" (Spiro 2009i). Merced uses an approval plan that all of the librarians help "fine-tune."

- **Outsource services that do not need to be performed locally.** Rather than charging local staff with responsibility for acquiring and processing monographs, UC Merced contracted with YBP Library Services to provide shelf-ready books that are already cataloged, labeled, and RFID-security-tagged. If library staff themselves had had to process the 30,000 volumes that were in the stacks when UC Merced opened, 25 people would have been needed. Instead, one librarian and one library assistant were able to carry out the work (Barclay 2007, 111). Since the library opened, the two-person technical/collection services team has put almost 85,000 books on the shelves. According to Miller, "It's a two-man show. It's all about specification and writing checks. We don't do piecework."

  UC Merced has also outsourced other operations that do not have to be performed locally. Instead of cataloging gift books in-house, it sends them to OCLC. As a result, Merced has no backlogs, and the quality of the catalog is "arguably better" than what could be done in-house because of the greater efficiencies of outsourcing (Spiro 2009i). UC Merced has also outsourced hosting the library Web site, getting access to sophisticated tools for a low cost without having to worry about maintaining servers. The CDL handles licensing e-resources. In a sense, Merced even outsources some operations to its users, who employ self-checkout machines to borrow materials. Miller claims that outsourcing has resulted in efficiency without sacrificing quality: "We still use library professionals to select and catalog books and manage databases, but they happen to be distributed everywhere, they're not in our building. … The product is what we care about."
- **Leverage collaborations.** UC Merced is one of 10 universities in the UC system, so it has access to more than 34 million volumes collectively held by UC universities through in-house, patron-initiated interlibrary loan. UC Merced likewise benefits from being affiliated with the CDL, an eleventh, virtual campus in the UC system. CDL aggressively negotiates license agreements, getting the best deals for its members. It also coordinates programs such as Google Books scanning and membership in the HathiTrust, which enables members of the UC system to search a corpus that includes more than 4 million volumes digitized by member institutions. As a result, Miller says, "Here at Merced we are incredibly leveraged and have access to incredible resources that another start-from-scratch library wouldn't have." The UC Merced Library has also pursued collaborations to digitize unique collections, such as works of Japanese art from

the Ruth and Sherman Lee Institute for Japanese Art at the Clark Center in Hanford, California.

- **Favor electronic over print journals.** Instead of providing print journals, UC Merced offers access to approximately 20,000 full-text journals through CDL. As a result, "The strategic value of not having print periodicals is that our book stacks will be filled with books instead of long runs of unused bound journals" (Barclay 2007, 111-112). (Merced does offer a small browsing collection of about 100 popular print magazines selected by students.)
- **Rethink the reference model.** The UC Merced Library has no reference desk. The decision to not offer a reference desk was driven partly by limitations in staffing, but primarily by the belief that it is not the most efficient way to provide service because it reaches only people who come into the library building. Instead, the library offers reference service through phone, e-mail, and chat. It also uses digital signage to communicate important information to users. Instead of adopting the "reference librarian" label, UC Merced employs the title "user communication and instruction librarian." Contrasting providing assistance to students via a mobile phone with traditional reference service, former UC Merced Librarian Michelle Jacobs comments, "I have reached almost twice as many students as when I sat at a reference desk. I've had time to explore new and innovative things and get a grasp on what makes the latest generation work. They like this technology, and who am I to tell them that this is not the best way to communicate?" (quoted by Carlson 2007). In the long term, UC Merced hopes to scatter librarians throughout the departments, so that librarians go to users rather than vice versa. According to Miller, the goal is to integrate librarians with the university: "the better we do our jobs, the less visible we are."
- **Deliver instruction flexibly.** Librarians recognized that as enrollments at Merced increased, delivering in-person library orientations would not be scalable. Instead of giving the same library tour and instruction in how to use the catalog and other library resources over and over, Merced loaded programmed tours onto 15 iPod Touches that users can check out from the library. The interactive tour guides students through the library and teaches and tests learning along the way. The video was shot by student assistants.
- **Make the physical space flexible.** The physical UC Merced Library is located in the Kolligian Library building and is designed to be "flexible enough to serve a variety of emerging, somewhat unpredictable needs" (UC Merced Library 2009b). Deputy University Librarian Donald Barclay compares the library space to public lands, so that the same space may be used for quiet study, collaborative work, or socializing. Wireless is available throughout the library, which features a mix of casual and formal spaces, such as a traditional reading room.
- **Abandon print reserves.** Merced does not have print reserves, which Miller says allow patrons to monopolize resources. Instead,

it offers online reserves integrated into the course-management system.
- **Evolve with technologies.** Merced does not have public desktop computers, but it does check out laptops to students who do not want to bring their own to the library. In the 2008–2009 academic year, 2,700 users checked out laptops 66,000 times. Over time, though, Miller anticipates that mobile devices will become prevalent and that the library will no longer have to provide laptops.

According to Miller, many libraries will be, or already are, pursuing approaches similar to UC Merced's. Merced represents a "next-step evolution of what all libraries will be doing" (Spiro 2009i). Miller predicts that libraries increasingly will operate at the network level: "I think that the generic stuff that libraries do is going to end up being handled at a very high networked level, a much more centralized assembly-line environment. What's unique to the campus will be dealt with on individual campuses."

3.2 CSU–Channel Islands

CSU–Channel Islands (CSUCI) is a four-year public university that is the newest in the 23-campus California State University system. Spun off from the Ventura branch of California State University–Northridge, it is located in Ventura County on the grounds of the former Camarillo State Mental Hospital. It opened in 2002 and graduated its first freshman class in 2007. When Head of Collections and Technical Services Steve Stratton was hired, the library already had a small collection from the CSU–Northridge campus, as well as part-time staff hired by that campus. Although academic libraries that had been founded a little earlier had difficulty in achieving their goal of "going digital," CSUCI decided that "the timing was right" for building the library around digital collections (Spiro 2009c). In addition to providing access to 400,000 to 500,000 electronic titles (including through EEBO and ebrary), CSUCI has a core collection of about 75,000 print books, most of which were published in the past 10 years. It borrows a number of books, particularly older ones, through interlibrary loan; indeed, Stratton says that if you are starting up an academic library, "you better have a good consortium" (Spiro 2009c). The library has only a database and electronic products budget, not a book budget; materials purchases are funded by midyear or university foundation funds. CSUCI also provides full-text access to 11,000 e-journals and 22,000 journals indexed in various databases. The library offers a small browsing collection of 40 popular magazines and newspapers, but it does not check them in, and after four months they are taken to the gym for people to read there.

Stratton says that the library has been able to avoid a lot of problems through its reliance on electronic journals. Instead of having a "squadron" of acquisitions and technical services staff, CSUCI has one copy cataloger and one person who works half-time on acquisitions, half-time on interlibrary loan. Since CSUCI does not load

serials into the Online Public Access Catalog (OPAC), staff members do not have to continually update serials records. Likewise, since CSUCI acquires only a few print periodicals, staff members have much less processing work to do. Every staff member assumes multiple roles, working at both the circulation and the reference desks. As a result, librarians are aware of what faculty and students are looking for and what colleagues are working on. As Stratton comments, "The biggest benefit is that everyone is focused on providing service to the students, faculty, and the public" (Spiro 2009c). Although the format of the information to which CSUCI provides access may be primarily digital, the library's key goals—"providing service, getting people to material they need, instructing them how to find it and use it"—are the same as those of most academic libraries. In addition to purchasing or subscribing to electronic resources, CSUCI is creating its own digital collections, focusing on the local community. For example, it is working with local environmental groups to bring their materials into its institutional repository.

Even though it delivers many resources electronically, the CSUCI Library has a strong physical presence. Until 2008, the library was housed in an old hospital building. In spring 2008, CSUCI opened the John Spoor Broome Library, an award-winning, 137,000-square-foot facility designed by Lord Norman Foster. The design aims to "marry" the old and the new; parts of the old main hospital building were removed and a glass-and-steel structure was installed. The book stacks are viewable from any vantage point, and the library has the capacity for 250,000 to 300,000 bound volumes and 1,800 concurrent users (Bustler 2009). In addition to the library, the building houses the University Writing Center, a Learning Resource Center, and the Information Technology and Disability Resource programs (CSU–Channel Islands 2008). The CSUCI Library takes into account current research practices and preferences, particularly those of students, who expect everything to be available digitally. As Stratton says, "It's hard for me to imagine that the students would have the patience or interest in going off to the stacks to track down the paper journal that is 20 years old." He says that faculty have been "very supportive" of the approach the library has been taking, and that they like being able to find and access what they need. Likewise, students have been filling the building when school is in session and "love everything being digital." As Stratton comments, "It's all about service. I've not found anything on the downside to being digital" (Spiro 2009c).

3.3 Olin College of Engineering

In starting up the library at the Olin College of Engineering in Needham, Massachusetts, Director Dee Magnoni aimed to support the college's mission of facilitating entrepreneurship, innovation, and hands-on education. Olin, a private engineering college that was founded in 1997, has about 300 students; its first full freshman class arrived in 2002. Launching a new library meant that Magnoni had

the freedom to set the budget according to current goals, rather than having to work with an established budget and pull money away from traditionally funded areas. Magnoni believes that "students need to interact with information—sometimes electronically, but not always" (Spiro 2009o). To support hands-on learning, she built a realia collection, providing access to objects such as Legos, modeling clay, and crystal lattice sets. Starting up a library also enabled Magnoni to configure the space appropriately. Since Olin was wireless from the start, the library has only a few public workstations; most students bring their laptops. As a result, more space can be devoted to collaborative study areas.

Olin has faced challenges because of its small staff and lean budget. Three librarians staff the library, including the director, who also does reference and collection development; a technical services person, who handles the journals and database back ends; and a circulation/interlibrary loan/reserves person. The library has outsourced some work that could not be done internally, such as development of its Web site. The library has had to contend with budget cuts each year, which has resulted in reductions in its book and travel budgets. However, Olin is able to collaborate with other area libraries, including through the Babson Brandeis Olin Wellesley (B-BOW) consortium, to provide access to collections and services.

On the basis of her experience establishing the Olin Library, Magnoni recommends that new libraries consult with people who have a wide range of expertise. She set up an external advisory board, consisting of academics, people in industry, and vendors, that has helped shape the direction of the library. Magnoni also emphasizes the importance of focusing on the library's mission and asking questions about what supports that mission. For instance, "Do users need this service to be done here, or can it be outsourced? Does this work well electronically, or do they need it in tactile form? What makes sense now?" (Spiro 2009o). With the exception of libraries serving virtual universities, Magnoni believes academic libraries will probably need some physical collections, since not every type of information resource (such as design books and realia) is well suited for electronic delivery.

3.4 Ikeda Library at Soka University

Work on the Ikeda Library at Soka University, the oldest among the libraries profiled in this report, began in 2000, a year before the Orange County, California, private liberal arts college opened to undergraduates. When the library was being developed, librarians took a traditional approach to collections, focusing on print. Yet Saeed FakhriRavari, interim director of the Ikeda Library, doubts that many of these print works, particularly reference materials, are now being used. Since 2000, libraries have shifted toward electronic collections, as has the Ikeda Library, which aims for 100 percent of its reference collection to be digital. Given how frequently reference works are updated, FahkriRavari thinks that it does not make sense to purchase print editions, especially with the added expenses of shelving,

reshelving, heating, cooling, and so forth. Likewise, the Ikeda Library favors electronic journals, offering access to 16,000 full-text electronic periodicals included in 125 databases as opposed to 350 print journals (Ikeda Library, Soka University 2009). However, FahkriRavari believes that library collections should hold both print and electronic books, since "there are some materials that may be of value to have in print" (Spiro 2009p). As a new library, Ikeda Library has been able to use some technologies that established libraries might have difficulty adopting. For instance, it uses RFID tags to manage its collections, whereas many traditional libraries still use bar codes.

3.5 Health Science Libraries

Special libraries that serve the needs of specific disciplines have been quietly transitioning to primarily digital environments. Law, medicine, geology, and other special libraries have largely made the transition to mostly digital assets, in part because they rely less on monographs than do libraries that serve humanities scholars; users of these special library collections often assume that everything they need for their research is already online. We have looked, in particular, at some health science libraries that have emerged in the past few years to understand their decisions regarding the acquisition of digital versus print collections. Since 2000, more than 20 health science libraries (including branch libraries) have been opened or are being planned in the United States, including libraries at Florida State University, Commonwealth Medical College, the University of Arizona College of Medicine-Phoenix, and the State University of New York Upstate Medical University (Doyle 2008). New medical libraries are often small; a survey by Deb Rand of ten health science libraries opening between 2007 and 2011 suggests that they typically have five or fewer professional librarians and four or fewer support staff, not including students (Rand 2009). Since many in the medical professions prefer to access information online and a critical mass of digital resources is available, these new libraries tend to focus on digital collections. However, medical libraries have struggled to provide access to electronic versions of medical textbooks, since many publishers have not yet worked out a business model for selling online access. Although medical libraries may be making the transition to digital information faster than research libraries are, both types of libraries are experiencing similar transformations. In summer 2003, Michael Kronenfeld visited four health science libraries to study the changes occurring there (Kronenfeld 2005). He found seven major trends that reflect the emergence of the virtual library: a shift to electronic collections; staff working outside of the physical space of the library on liaison activities; the library building being used as a communal space; the development of new Web sites and tools; the shift in document delivery from print to electronic; the rise of nontraditional information formats such as databases; and computer-based operations. Many of these changes are also taking place at academic libraries generally.

3.5.1 Arizona Health Sciences Library—Phoenix

The Arizona Health Sciences Library (AHSL)–Phoenix,[13] a branch of the Arizona Health Sciences Library, opened on the Phoenix Biomedical Campus in 2007. This campus and the College of Medicine–Phoenix are a partnership of the University of Arizona and Arizona State University. Two librarians serve about 40 faculty and 120 medical students, a number that will grow to about 120 students per class. Occupying 1,100–1,200 square feet, the library has about 100 linear feet of stacks holding a core print collection of about 400 volumes. The rest of the space is devoted to study and collaboration. If funding is approved, the library will move into a new education building and have about 22,000 square feet. Since AHSL–Phoenix is a branch library, AHSL–Tucson manages e-resources and processes all books. AHSL–Phoenix controls most of its own budget, but allocates part of its budget to Tucson to cover some cooperative resources. AHSL–Phoenix has no print journals, and most of its collection is available online. Indeed, the library bills itself as a primarily digital library: "In Phoenix you will notice a focus on the use of electronic materials, except in the case of books not yet available in digital format. What is exciting about our digital collections is the fact that our resources will be available to you wherever and whenever you need them, at your fingertips" (Arizona Health Sciences Library 2009). In establishing and running a new academic library, Head Librarian Jacque Doyle emphasizes the importance of knowing the target audience. Because the institution is so small, the librarians do a lot of outreach to students and faculty, and they carry cell phones so that they can be reached easily. The library functions as a social space, as medical students come in to study together, read textbooks, or use the refrigerator and microwave. When people enter the library for the first time, they may be surprised by its small size and lack of books, but "most people know that it's much bigger than it looks" (Spiro 2009j). According to Doyle, the library always planned to be primarily digital, an approach that faculty and students embrace. However, she has been surprised by the difficulty of purchasing network access to electronic versions of many of the medical textbooks they want, noting that publishers often will sell e-access only to individuals who purchase the print version. Although AHSL–Phoenix is a new library, Doyle suggests that it is not very different from most other health science libraries in its emphasis on digital collections, user service, and collaboration with other libraries.

3.5.2 Learning Resource Center, A. T. Still University of the Health Sciences

At the Learning Resource Center (LRC) of A. T. Still University of the Health Sciences in Mesa, Arizona, a staff of three serves a faculty of 70 to 80 and approximately 400 students. Although the library has a small print collection of 110 journals, it offers access to 2,400 electronic journals and to twice as many e-books as print books. When Michael Kronenfeld was hired in 2002 to start up the library, he set

[13] http://www.ahsl.arizona.edu/about/phoenix/.

out to identify its key strategic goal. "As long as you have a direction you're going … you find that you get there faster," he says. (Spiro 2009m). Kronenfeld decided to focus on evidence-based medicine, helping students integrate evidence into clinical practice so that they could improve the quality of care. Such a focus meant that the library added services that are not necessarily typical; for example, it built an academic computing-support service, helps faculty develop online content for courses, provides poster printing at no charge, and offers assistance in creating multimedia content. Kronenfeld works to support students in using electronic resources effectively and coteaches classes on evidence-based medicine. For clinicians, access to electronic resources is critical, since they typically do not have time to come to a library to consult print books and the only print resources they have on hand are textbooks, which rapidly become outdated. Kronenfeld says the emphasis on evidence-based medicine reflects changes in how libraries deliver information. In the 1990s, libraries focused on providing access to as much full-text information as possible; today, they must figure out how to make this "vast array of information useful to end users who have grown up thinking search is Google" (Spiro 2009m). Toward this end, the LRC not only helps faculty integrate evidence-based medicine into the curriculum but also tries to make it easier for them to discover relevant information by developing a federated search engine for its collections and working on a clinical search tool. To ensure that clinicians can practice evidence-based medicine after they graduate, the LRC offers alumni access to some of its resources through the alumni library and provides free document delivery for up to three resources per month.

Collaboration has been important to LRC's success. When Kronenfeld arrived, he identified tasks that could be accomplished immediately, such as joining a statewide consortium to provide access to library resources, which he characterizes as a "quick win" (Spiro 2009m). Collaboration enables the LRC to get more resources for less money and to share ideas.

Kronenfeld believes that the library's future depends on how open the library profession is to change: "Our job is to keep up and keep moving the library toward supporting its and the university's strategic goals." Kronenfeld has witnessed significant changes in how medical libraries work. From the 1970s to the mid-1990s, librarians performed searches for patrons using resources such as MedLine. With the emergence of the Web, users could run their own searches, changing the role of the librarian but not the key goal of facilitating access to evidence. Now librarians' roles are shifting again, as they create information systems and train people to use them.

3.6 International Libraries: NYU Abu Dhabi

In an academic building boom, U.S. colleges and universities such as Cornell, Texas A&M, Georgetown, Massachusetts Institute of Technology, and Michigan State are creating campuses in Dubai, Abu Dhabi, and Qatar (Krieger 2008). In 2007, NYU announced that it

would launch a "comprehensive liberal arts campus" in Abu Dhabi, providing labs, dorms, and a library and information technology facilities. NYUAD opened a small library on its downtown campus in January 2010, and the library for the main campus is expected to open in about four years. Initially, the library will be 99 percent digital; ultimately, it will be approximately 95 percent digital. Library Director Steve Shorb says that 5 percent of the collection will need to be physical, since for some types of information books work better visually and ergonomically. Nonetheless, the library will never have more than a few hundred print periodicals, reflecting the general movement toward electronic journals. According to Shorb, "We'll be building up from zero to 250 print subscriptions at same time that NYU is ramping down to far fewer print subscriptions than they have now" (Spiro 2009f).

Although the fundamentals of the library will stay the same, NYUAD is rethinking them for a digital context, billing itself as "a library for the 21st century" (NYU Abu Dhabi Library 2010). It will provide access not only to content that it would be difficult for individuals to access on their own but also to tools for analyzing and organizing that information. Shorb plans to expand the range of material the library typically holds and the types of services it provides, emphasizing data manipulation and visualization as well as the digitization and delivery of multimedia resources. Because NYUAD will rely on the main campus of NYU for access to many library resources, it plans to develop new ways of rapidly delivering those resources, including on-demand digitization of resources, print on demand, and physical delivery. The government of Abu Dhabi will pay for the digitization of newly created collections along with selected published and unpublished print resources in NYU's Bobst Library to support research and teaching. NYU Librarian Carol Mandel calls this approach "digital collection development" (Oder 2009b). NYUAD will not only provide access to scholarly information but will also collaborate to disseminate it. NYUAD will partner with the NYU Press to publish scholarly resources produced at the university, such as scholarship focused on the Persian Gulf, digital scholarship, and faculty lectures. In order to achieve its vision, NYUAD will recruit generalist librarians with strong technical skills and a service orientation. Even as it provides digital access to many resources, NYUAD will also focus on the library as place. Shorb aims for the physical library to be a "third place," a social space for "exchanging knowledge and learning in different environments," so it will have "turbo-charged" group study rooms as well as other spaces focused on collaboration (Spiro 2009f). The entire library will function as a learning commons.

4. Recommendations

All seven libraries profiled above have recently faced the question we posed at the beginning of this report: How should a new university approach establishing its library? Although the new academic

libraries discussed here serve different constituencies and have different areas of focus, some common recommendations for startup libraries emerge from our conversations with their leaders.

1. Define the core mission.
In establishing new libraries, leaders have oriented themselves around their key goals. As Bruce Miller says, new libraries can pose new questions and challenge traditional assumptions. Instead of asking, "Where should the reference desk go," UC Merced asked, "What is the purpose of reference? How do we fulfill that purpose?" Michael Kronenfeld of A. T. Still suggests that a library should use its strategic goals to guide its direction. Likewise, Dee Magnoni of Olin emphasizes that new libraries should focus on what supports the library's goals and ask, "Can this service be outsourced? Does this work well electronically, or do users need it in tactile form? What makes sense now?" (Spiro 2009o).

2. Be flexible in offering access to information resources.
Print and digital resources currently meet different purposes, with print supporting immersion and long-form reading, and digital supporting searching and immediate access. Many libraries founded after 2000 provide access primarily to electronic journals, retaining a small browsing collection of popular magazines and core journals. Such a decision frees space for uses other than storing back issues, staff for purposes other than processing and shelving journals, and funds for priorities other than offering both print and electronic journals. Nonetheless, new academic libraries have not really replaced print books with electronic books. As we have demonstrated, a significant number of books are not available electronically. In addition, libraries find that print remains the appropriate format for certain types of resources, including art and design books. Yet these startup libraries do have smaller print collections than the typical academic library, partly because they do not have large legacy collections, and partly because they chose to focus on collecting books directly relevant to teaching and research priorities rather than adopting a "just in case" approach to collection development. They also rely heavily on cooperative relationships to borrow books not in their collections.

With the shift to electronic, libraries are no longer limited to what they have in their local collections: they can provide rapid access to whatever researchers need. As Price and McDonald say, "The shift to digital represents another step toward an end to the limitation of your local collection. It used to be that you were limited to what your library had. Then the union catalog and ILL [interlibrary loan] broadened what books you could get. Now, if you can broaden to whatever is available and are not limited to a local collection, you can get whatever you want" (Spiro 2009a). Shifting to e-books enables libraries to adopt a patron-driven acquisition model, whereby patrons request what they need and can have it downloaded almost instantaneously. Libraries set up restrictions to ensure that one person does not consume too much of the budget and that appropriate

books are purchased. With this model, patrons feel empowered and part of the library. As Polanka says, "You give the patrons what they want, when they need it" (Spiro 2009g). Items purchased through patron-driven acquisition appear to be used more frequently than books purchased by librarians—not only by the patron who made the initial selection but also by other users. In a 2009 study, Price and McDonald used data provided by EBL to compare usage patterns for user-selected ("patron-acquired") and preselected ("librarian-acquired") e-books at five libraries that supported both acquisition models. They found that user-selected e-books were used more frequently and by more unique users; further, user-selected collections did not reduce breadth in the subject areas collected (Price and McDonald 2009). This flexible approach not only makes economic sense but also serves the ultimate goal of the library to support research: "If we shift to buying what people want at the university, we pay only after it's been used" (Spiro 2009a).

3. Collaborate with other institutions to share resources and knowledge. As Ross Atkinson argues, libraries face many complex challenges, such as converting print to electronic formats, developing collection strategies, negotiating with publishers on purchasing, and preservation. They can meet these challenges only through the often-difficult work of cooperating, sharing resources and responsibilities, and leveraging collective power (Atkinson 2005). Libraries are operating in a networked world where resources can be shared regardless of location and where scholarship is conducted in a rapid, distributed, and collaborative fashion (Schonfeld and Housewright 2008, 31-32). Libraries already cooperate through cataloging, interlibrary lending, consortial purchases, collaborative software development, joint research projects, and other activities. Now they are beginning to cooperate on sharing print and digital collections (Spiro 2009h). New libraries do not have legacy collections, which means that they often must depend on other libraries for access to older or more unique works, as well as purchase back files. If a library were to decide to provide access only to digital collections, it would need to make business arrangements so that it could rapidly get access to print resources requested by researchers. At UC Merced, for example, the library can support the research needs of faculty and students because it is a member of the UC system and has access to all the resources collectively held by its 10 campuses. If UC Merced could not leverage the collective resources of the UC system, Miller suggests, it would have difficulty meeting its mission as a research library. Likewise, CSU–Channel Islands and CSU–Monterey Bay depend on the central CSU office to negotiate licenses for electronic resources and on the other CSU campuses to deliver resources via interlibrary loan. Indeed, most of the new libraries we examined belonged to a consortium or were a branch of an established university or university system.

Several libraries are exploring a collaborative, distributed model for library collections, whereby collections are shared across a

network of libraries. Constance Malpas suggests that libraries need to rethink how they provide access to collections and to move toward a cooperative model: "You need collections, you need access to material that is available in online formats, but the way you do that is not the way we used to, where everyone buys their own copy." Shifting from local collections to a network model will require establishing efficient workflows. How much replication is necessary to ensure rapid access to print and digital resources now and preservation in the long term? Liberal arts colleges and universities such as the CONSORT colleges and Tri Colleges are implementing cooperative collection-development plans so that they do not have to acquire multiple copies of the same work (Luther et al. 2003). Faced with a severe space crunch and high costs for real estate, NYU is partnering with OCLC Research to explore how much of its physical inventory it needs to retain to meet user needs and its preservation responsibilities (Spiro 2009h). NYU is exploring the possibility of reducing locally managed physical collections in favor of cooperative agreements with large-scale shared print and digital repositories, withdrawing low-use materials that can be more cost-effectively acquired from other suppliers. The goal is to optimize the physical inventory in view of changing preservation and access requirements, selectively externalizing collection-management functions that no longer deliver distinctive local value. A quantitative and qualitative assessment of the duplication between NYU library holdings and large preservation repositories will provide insights into the potential scope and scale of this externalization (Spiro 2009h).

HATHI Trust and ReCAP are partnering with NYU in this project to help characterize the service requirements and business models needed to support a large-scale shift from locally managed inventory to "cloud-sourced" research collections. We can expect other research-intensive academic libraries to explore this approach in the years to come. Library support for the academic mission of the university is no longer measured by the size of the local book collection, but by participation in collective preservation and access strategies. College and university libraries need to be aware of "where they fit in the larger research ecology" and to assume collective responsibility for maintaining the health of the system. In the long term, this will likely entail a strategic reallocation of resources away from locally managed print collections and toward cooperatively managed digital aggregations (Spiro 2009h).

4. Develop new service models.

With the shift to electronic resources, libraries can focus more on providing services than on managing collections, including support for manipulating and disseminating digital information. For instance, they can replace the reference desk with a distributed model, as UC Merced has done, and they can work closely with faculty on information-fluency or research efforts, as A. T. Still has done. When asked to predict what the shift to digital collections will mean for libraries, Price and McDonald suggest, "Libraries will be

services-based organizations and not collections-based any more. … Libraries should be helping users create, manage, and manipulate information, and not just acquiring it and making it available to their constituency" (Spiro 2009a). One emerging role for librarians is providing research support to scholars, particularly (though not exclusively) in the digital arena. Many humanities researchers are hungry for assistance from librarians on their research projects, viewing them as important team members who could help locate, describe, organize, make available, and preserve diverse forms of data (Spiro 2009l). Librarians may also find a role in helping digitize, migrate/maintain, and make available the research notes and data collected by scholars in the course of their work. In archaeology, for example, field notes and photographs—typically the only complete record of a site—are often relegated to file cabinets after a dig is complete; librarians can assist in archiving these materials and making them more widely available for use by other researchers (Spiro 2009l).

5. Reimagine librarian roles.
Rather than assigning librarians to fixed roles, many of the new libraries we examined require that their staff take on a variety of functions. In part, this approach was driven by the small size of the staff, but it also reflects a holistic understanding of the functions of the library and a sense that librarians are more engaged if they participate in a variety of activities. At several of these libraries, processing and cataloging materials is done elsewhere, enabling librarians to focus more on providing user services. New libraries have the advantage of being able to recruit staff with the desire and experience to take on this more flexible role.

6. Build flexible facilities that support collaboration and interaction.
Even as many new libraries embrace digital collections (particularly journals), they continue to emphasize the importance of the physical space of the library in supporting learning, reflection, and collaboration. If a library did not have to devote much of its space to print collections, how would the space be configured? As Sue Polanka suggests, "I would see the library taking on an entirely different look and feel—no longer rows of bookshelves, but an expanded information commons" (Spiro 2009g). Indeed, many new libraries do embrace the information commons model, reflecting the general trend in library building and renovation projects. Several of the libraries we profiled recently opened flexible, attractive new facilities that feature collaborative work spaces and enable users to move furniture around, share large computer displays, and retreat to quiet areas. Some of these facilities are designed so that they can be easily repurposed as new needs arise. Yet libraries at public universities are also somewhat constrained in how they design space, with the requirement that they have a certain amount of shelving and can seat a minimum percentage of the total student population. Other new libraries are planning similar facilities, reflecting a belief that the physical library remains relevant in the digital age. But one of our interviewees

questioned whether the investment in library as social space was merited, given libraries' limited resources. Is providing study and collaborative space core to the mission of the library? Can other organizations on campus fulfill that role? What unique expertise does the library possess that it can bring to the research enterprise?

7. Plan for preservation.
As libraries shift into new modes of providing access to information, there will inevitably be a change in the culture and practice of research. Even as libraries deaccession print materials, at least some copies of these materials should be preserved. Some researchers need to study the physical properties of books, such as bindings, paper, type, and illustrations. Others do not trust that digital resources will be around for the long term and want a print backup. Yet it is probably not necessary for every library to retain its own copy of a print book. Libraries can implement print preservation cooperatively, perhaps by establishing regional networks of shared print repositories (Reilly 2003). Sharing print collections will require establishing a workflow for managing these materials, developing trusting partnerships with other institutions, working through legal prohibitions about disposing of print, and ensuring that sufficient copies are kept and that the digital copy is of sufficient quality (Malpas and Massie 2009). Likewise, libraries, scholars, and publishers must collaborate in developing workable, trustworthy models for long-term preservation of digital resources.

5. Future Work

This is an exploratory study to identify some of the challenges that new academic libraries will face, as well as to examine the experiences of several recently established libraries as they continue to move into the digital era. We believe more research needs to be done on the following topics:

1) *Electronic books.* Although much research has been done on the implications of e-journals for libraries and research practices, the significance of electronic monographs has not been studied to the same extent, perhaps because they have not yet reached the tipping point. We need to know more about economic models for electronic books, plans for long-term preservation and access, how to make e-books usable for students and researchers, how to negotiate rights concerns, and more.
2) *New academic libraries.* Our study is just an initial step. We would like to see a more systematic study that goes beyond interviews with library leaders and includes surveys and focus groups with staff and users. In addition, we believe that the scope should be expanded to include new universities not in the United States.
3) *Preservation*. Much research is already being done on preserving access to scholarly materials, but libraries and researchers need to feel confident that workable solutions are being implemented before they are comfortable with the transition to digital.

Initiatives such as Portico and LOCKSS have only recently begun to tackle the challenges of preserving e-books, and much more work remains to be done on that front. Further, economic, organizational, and technical models for preserving content need to be developed.

6. Conclusion

This is a time of rapid change for libraries. That change is visible at new academic libraries, which are transitioning toward digital collections but still providing access to core print collections. Given technical, cultural, policy, and economic obstacles, it is probably premature for most libraries to decide to provide access only to electronic collections, particularly when it comes to monographs. Yet as new libraries have found with their electronic journal collections, shifting away from print can be liberating, freeing staff and space for library services such as research support. Libraries need to be realistic about the obstacles facing the transition to digital collections but also future oriented in preparing for such a shift. It will be important for the research library to continue to demonstrate its perceived value by the faculty, students, and researchers it serves if it is to survive as a core part of the institution to which it belongs. For researchers in fields that have already largely migrated to predominantly digital formats, the library is often viewed as having a diminished role. As the trend toward an all-digital environment continues, the services and support that libraries provide must address users' needs in ways that distinguish them from other commercially available services. Libraries are recognizing this in varying degrees, but the continual shift is undeniable. Whereas it would have been difficult in 1995 for an academic library to be all-digital, in 2010 the primarily digital academic library seems to be on the horizon. As Daniel Greenstein suggests, "The library that acts as a steward will have to learn what it means to capture and persistently manage new vehicles of information. It will have to change in order to stay the same" (quoted by Olsen 2005).

Works Cited

Note: URLs are current as of May 15, 2010.

Abbott, Andrew, James Chandler, Martin Feder, Judith Nadler, Richard Rosengarten, Martha Roth, and James Vaughan. 2006. *Provost's Task Force on the Library—Final Report*. Chicago: University of Chicago. Available at http://www.lib.uchicago.edu/e/about/finalreport.html.

Anderson, K. J., R. S. Freeman, J. P. V. M. Hérubel, L. J. Mykytiuk, J. M. Nixon, and S. M. Ward. 2002. Buy, Don't Borrow: Bibliographers' Analysis of Academic Library Collection Development through Interlibrary Loan Requests. *Collection Management* 27(3/4): 1–12.

Anson, Catherine, and Ruth R. Connell. 2009. E-book Collections. SPEC Kit, 313. Washington, DC: Association of Research Libraries. Available at http://www.arl.org/news/pr/spec313.shtml.

Arizona Health Sciences Library. 2009. About AHSL Phoenix. Available at http://www.ahsl.arizona.edu/about/phoenix/.

ARL (Association of Research Libraries). 2009. LibQUAL+®. Washington, DC: ARL. Available at http://www.libqual.org/home.

ARL, American Library Association, and Association of Colleges and Research Libraries. 2009. Library Associations Letter to U.S. Department of Justice, December 15. Available at http://www.arl.org/bm~doc/antitrustdivasa-final.pdf.

Association of American Publishers (AAP). 2010. "AAP Reports Book Sales Estimated at $23.9 Billion in 2009," news release, April 7. Available at http://www.publishers.org/main/PressCenter/Archicves/2010_April/BookSalesEstimatedat23.9Billionin2009.htm.

Atkinson, Ross. 2005. Six Key Challenges for the Future of Collection Development. Presented at the Janus Conference on Research Library Collections, Ithaca, NY, October 10. Available at http://ecommons.cornell.edu/handle/1813/2608.

Bajarin, Tim. 2009. Can the Apple Tablet Kill the Kindle? *PC Magazine*, October 26. Available at http://www.pcmag.com/article2/0,2817,2354658,00.asp.

Baker, David, and Wendy Evans. 2009. *Digital Library Economics: An Academic Perspective*. Oxford: Chandos.

Baker, Nicholson. 2009. Kindle and the Future of Reading. *The New Yorker*, August 3. Available at http://www.newyorker.com/reporting/2009/08/03/090803fa_fact_baker.

Ball, Rafael. 2009. E-books in Practice: The Librarian's Perspective. *Learned Publishing* 21(1): 18–22.

Band, Jonathan. 2009. *A Guide for the Perplexed: Libraries and the Google Library Project Settlement*. American Library Association, Association of Research Libraries, and Association of College and Research Libraries. Available at http://wo.ala.org/gbs/alaarl-summary-document/.

Bar-Ilan, Judit, and Noa Fink. 2005. Preference for Electronic Format of Scientific Journals—A Case Study of the Science Library Users at the Hebrew University. *Library & Information Science Research* 27(3): 363–376. doi:10.1016/j.lisr.2005.04.011.

Barclay, Donald A. 2007. Creating an Academic Library for the Twenty-First Century. *New Directions for Higher Education* 2007(139): 103–115.

Biglione, Kirk. 2007. DRM for Books: Will Publishers Learn Anything from the Music Industry's Mistakes? *Medialoper*, June 25. Available at http://medialoper.com/hot-topics/print/drm-for-books-will-publishers-learn-anything-from-the-music-industrys-mistakes/.

Borgman, C. L., G. C. Bowker, T. A. Finholt, and J. C. Wallis. 2009. Towards a Virtual Organization for Data Cyberinfrastructure. In *Proceedings of the 2009 Joint International Conference on Digital Libraries*, 353–356. New York: Association for Computing Machinery.

British Library. 2005. "British Library Predicts 'Switch to Digital by 2020,'" news release, June 29. Available at http://www.bl.uk/news/2005/pressrelease20050629.html.

———. 2008. The British Library's Strategy 2008–2011. Available at http://www.bl.uk/aboutus/stratpolprog/strategy0811/index.html.

Brown, Laura, Rebecca Griffiths, and Matthew Rascoff. 2007. *University Publishing in a Digital Age*. ITHAKA report, July 26. Available at http://www.ithaka.org/saved-pages/strategic-services/university-publishing.

Buckman, Rebecca. 2006. Rice University Revives Its Press in Digital Model. *Wall Street Journal*, July 13, sec. Technology. Available at http://online.wsj.com/public/article/SB115275720041205358-86er1J1h9klBT9EPKSlQztt8HXA_20060812.html?mod=tff_main_tff_top.

Bustler. 2009. CSU Channel Islands Library Wins Design Award. Posted January 5. Available at http://www.bustler.net/index.php/article/csu_channel_islands_library_wins_design_award/.

C-LIB Subcommittee on Digital Information Technologies in the Research Library Environment at Stanford. 2008. *Report of the C-LIB Subcommittee on Digital Information Technologies in the Research Library Environment at Stanford* (8 September). Palo Alto, CA: Stanford University. Available at http://facultysenate.stanford.edu/2008_2009/reports/SenD6153_c_lib_dig_info.pdf.

Caplan, Priscilla. 2009. *Understanding PREMIS*. Washington, DC: Library of Congress. Available at www.loc.gov/standards/premis/understanding-premis.pdf.

CARL Copyright Committee Task Group on E-Books. 2008. E-Books in Research Libraries: Issues of Access and Use. Available at http://www.carl-abrc.ca/projects/copyright/pdf/CARL%20E-Book%20Report-e.doc.

Carlson, Scott. 2002. Do Libraries Really Need Books? *Chronicle of Higher Education*, July 12. Available at http://chronicle.com/weekly/v48/i44/44a03101.htm.

———. 2005. Shelve under 'E' for Electronic. *Chronicle of Higher Education*, April 4. Available at http://chronicle.com/article/Shelve-Under-E-for/65383/.

———. 2007. Are Reference Desks Dying Out? *Chronicle of Higher Education*, April 20. Available at http://chronicle.com/article/Are-Reference-Desks-Dying-Out-/20976/.

CASPAR Project. 2009. CASPAR Digital Preservation User Community. Available at http://www.casparpreserves.eu/.

Center for Research Libraries. Certification and Assessment of Digital Repositories. Available at http://www.crl.edu/archiving-preservation/digital-archives/certification-and-assessment-digital-repositories.

Cleto, Cynthia. 2008. ATG Special Report—10 Steps to Implementing an eBook Collection: A Guide for Librarians. *Against the Grain* (February): 47–48.

CLOCKSS. 2008. CLOCKSS: A Trusted Community-Governed Archive. Available at http://www.clockss.org/clockss/Home.

Conley, J. P, and M. Wooders. 2009. But What Have You Done for Me Lately?: Commercial Publishing, Scholarly Communication, and Open-Access. *Economic Analysis and Policy* 39(1): 71.

Connaway, Lynn, and Stephen R. Lawrence. 2003. Comparing Library Resource Allocations for the Paper and the Digital Library. *D-Lib Magazine* 9(12). doi:10.1045/december2003-connaway. Available at http://www.dlib.org/dlib/december03/connaway/12connaway.html.

Connaway, Lynn, and Heather Wicht. 2007. What Happened to the E-book Revolution?: The Gradual Integration of E-books into Academic Libraries. *Journal of Electronic Publishing* 10(3).

Cornell University Library. arXiv.org e-Print archive. Available at http://arxiv.org/.

Corson-Finnerty, Adam. 2009. The Great Library in the Sky Version 0.6. Lulu. Available at http://www.lulu.com/content/paperback-book/the-great-library-in-the-sky-version-06/7190627.

COUNTER. COUNTER—Online Usage of Electronic Resources. Available at http://www.projectcounter.org/index.html.

Courant, Paul N. 2006. Scholarship and Academic Libraries (and Their Kin) in the World of Google. *First Monday* 11(8). Available at http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/1382/1300.

———. 2008. The Future of the Library in the Research University. In *No Brief Candle: Reconceiving Research Libraries for the 21st Century*, 21–27. Washington, DC: Council on Library and Information Resources. Available at http://www.clir.org/pubs/abstract/pub142abst.html.

Courant, Paul N., and Matthew "Buzzy" Nielsen. 2010. *On the Cost of Keeping a Book*. In *The Idea of Order: Transforming Research Collections for 21st Century Scholarship*, 81–105. Washington, DC: Council on Library and Information Resources. Available at http://www.clir.org/pubs/abstract/pub147abst.html.

Cox, John. 2008. Making Sense of E-book Usage Data. *The Acquisitions Librarian* 19(3/4). Available at www.library.nuigalway.ie/coxe-books.pdf.

CSU–Channel Islands. 2008. John Spoor Broome Library Construction Facts. Available at http://library.csuci.edu/info/libraryarchitecture.html.

CSUMB (CSU–Monterey Bay) Library. 2009. The Library as Place. Available at http://web.archive.org/web/20060910055109/http://csumb.edu/site/x5762.xml.

Cyberinfrastructure Council. 2007. *Cyberinfrastructure Vision for 21st-Century Discovery*. Arlington, VA: National Science Foundation. Available at http://www.nsf.gov/pubs/2007/nsf0728/.

Darnton, Robert. 2008. The Library in the New Age. *The New York Review of Books* 55(10), June 12. Available at http://www.nybooks.com/articles/21514.

De Rosa, Cathy, Joanne Cantrell, Janet Hawk, and Alane Wilson. 2006. *College Students' Perceptions: Libraries & Information Resources*. Dublin, OH: Online Computer Library Center. Available at http://www.oclc.org/reports/perceptionscollege.htm.

Declerck, Luc, Amy Kautzman, Patti Martin, Terry Ryan, and John Riemer. 2005. *Rethinking How We Provide Bibliographic Services for the University of California*. The University of California Libraries. Available at http://libraries.universityofcalifornia.edu/sopag/BSTF/Final.pdf.

Dempsey, Lorcan. 2006. The (Digital) Library Environment: Ten Years After. *Ariadne* 46 (February). Available at http://www.ariadne.ac.uk/issue46/dempsey/intro.html.

Digital Curation Centre. 2010a. *What is the Digital Curation Centre?* Available at http://www.dcc.ac.uk/.

———. 2010b. Frequently Asked Questions about DRAMBORA. Digital Curation Centre. Available at http://www.dcc.ac.uk/digital-curation/digital-curation-faqs/drambora.

Dixon, Kim. 2005. A New Library Survey Supports Expanding Space for Knowledge. *The University of Chicago Chronicle*, November 3. Available at http://chronicle.uchicago.edu/051103/space-knowledge.shtml.

Doyle, Jacque. 2008. Survey, 2008, of NEEHSL Members' Plans. *NEEHSLs* and Beyond*. Blog available at http://newemergingmedicallibraries.wordpress.com/summary-of-the-survey-earlier-this-year.

DPE (Digital Preservation Europe). 2009. Digital Preservation Europe. Available at http://www.digitalpreservationeurope.eu/.

DuraSpace. 2009. DuraCloud Project. Available at http://duraspace.org/duracloud.php.

ebrary. 2007. *2007 Global Faculty E-book Survey*. Available at http://www.ebrary.com/corp/collateral/en/Survey/ebrary_faculty_survey_2007.pdf.

ePrints. ROARMAP. Available at http://www.eprints.org/openaccess/policysignup/.

Epstein, Jennifer. 2009. A Win for the Stacks. *Inside Higher Ed*, November 13. Available at http://www.insidehighered.com/news/2009/11/13/syracuse.

Farley, T., J. Broady-Preston, and T. Hayward. 1998. Academic Libraries, People and Change: A Case Study of the 1990s. *Library Management* 19(4): 238–251.

Fenton, E. G., and Roger Schonfeld. 2004. The Transitional Period of the Periodicals Format Shift. *LIBER Quarterly* 14: 3–4.

Ganapati, Priya. 2009. Why E-Books Are Stuck in a Black-and-White World. *Wired.com*, Gadget Lab, June 9. Available at http://www.wired.com/gadgetlab/2009/06/blackandwhite_ebooks/.

Glazer, Gwen. 2010. Cornell University Library Engages More Institutions in Supporting arXiv. *Cornell University Library News*, January 21. Available at http://news.library.cornell.edu/news/arxiv.

Grafton, Anthony. 2009a. Apocalypse in the Stacks? The Research Library in the Age of Google. *Daedalus* 138(1): 87–99.

———. 2009b. Google Books and the Judge. *The New Yorker*, September 18. Available at http://www.newyorker.com/online/blogs/books/2009/09/google-books-and-the-judge.html.

Grigson, Anna. 2009. Evaluating Business Models for E-Books Through Usage Data Analysis: A Case Study from the University of Westminster. *Journal of Electronic Resources Librarianship* 21(1): 62. doi:10.1080/19411260902858623.

Hafner, Katie. 1995. Wiring the Ivory Tower. *Newsweek*, January 30. Available at http://www.newsweek.com/id/107007.

Harley, Diane. 2008. *The University as Publisher: Summary of a Meeting Held at UC Berkeley on November 1, 2007*. Berkeley, CA: UC Berkeley Center for Studies in Higher Education. Available at http://cshe.berkeley.edu/publications/publications.php?id=295.

Harley, Diane, Sophia Kyzys Acord, Sarah Earl-Novell, Shannon Lawrence, and C. Judson King. 2010. *Assessing the Future Landscape of Scholarly Communication*. Berkeley, CA: UC Berkeley Center for Studies in Higher Education. Available at http://escholarship.org/uc/cshe_fsc.

Hathi Trust. 2010. Mission & Goals. Available at http://www.hathitrust.org/mission_goals.

Henry, Charles, and Kathlin Smith. 2010. Ghostlier Demarcations: Large-Scale Text Digitization Projects and Their Utility for Contemporary Humanities Scholarship. In *The Idea of Order: Transforming Research Collections for 21st Century Scholarship*, 106–115. Washington, DC: Council on Library and Information Resources.

Higher Education Consultancy Group. 2006. A Feasibility Study on the Acquisition of e-Books by HE Libraries and the Role of JISC. Available at http://www.jisc.ac.uk/media/documents/jisc_collections/ebooks%20final%20report%205%20oct.doc.

Howard, Jennifer. 2009a. Library Protesters to Ohio State U.: Digital's OK, but Save Our Books! *Chronicle of Higher Education Wired Campus*, May 13. Available at http://chronicle.com/blogPost/Library-Protesters-to-Ohio-/7161/.

———. 2009b. Research Librarians Discuss How to Sell Scholars on Open Access, and More. *Chronicle of Higher Education*, October 18. Available at http://chronicle.com/article/Research-Librarians-Discuss/48870/?sid=at&utm_source=at&utm_medium=en.

———. 2009c. Columbia and Cornell Libraries Announce "Radical" Partnership. *Chronicle of Higher Education Wired Campus*, October 28. Available at http://chronicle.com/blogPost/ColumbiaCornell-Librar/8627/.

Hunter, Karen. 2007. The End of Print Journals—(In)Frequently Asked Questions. *Journal of Library Administration* 46(2): 119. doi:10.1300/J111v46n02_09.

Ikeda Library. 2009. About Ikeda Library, Soka University. Available at http://ikedalibrary.soka.edu/lib_library.html.

International Digital Publishing Forum. 2009. Wholesale eBook Sales Statistics. Available at http://www.idpf.org/doc_library/industrystats.htm.

ITHAKA. 2005. Portico: A Digital Preservation and Electronic Archiving Service. Available at http://www.portico.org/.

ITHAKA S+R. Print Collections Decision-Support Tool. Available at http://www.ithaka.org/ithaka-s-r/research/what-to-withdraw/print-collections-decision-support-tool.

JISC (Joint International Systems Committee). 2009. *JISC National E-Books Observatory Project: Key Findings and Recommendations*. Available at http://www.jiscebooksproject.org/reports/finalreport.

Johnson, Richard K., and Judy Luther. 2007. *The E-only Tipping Point for Journals: What's Ahead in the Print-to-Electronic Transition Zone*. Washington, DC: Association of Research Libraries. Available at http://www.arl.org/bm~doc/Electronic_Transition.pdf.

Kaufman, Paula, and Sarah Barbara Watstein. 2008. Library Value (Return on Investment, ROI) and the Challenge of Placing a Value on Public Services. *Reference Services Review* 36(3): 226–231. doi:10.1108/00907320810895314.

Kaufman, Paula T., and Scott Walter. 2008. The Library as Strategic Investment: Results of the University of Illinois "Return on Investment" Study from the Director's Perspective. Presented at the LibraryConnect Seminar 2008, Tokyo, Japan, July 16, 2008. Available at http://hdl.handle.net/2142/8768.

Kirchhoff, Amy J. 2009. Expanding the Preservation Network: Lessons from Portico. *Library Trends* 57(3): 476–489.

Kolowich, Steve. 2009. Libraries of the Future. *Inside Higher Ed*, September 24. Available at http://www.insidehighered.com/news/2009/09/24/libraries.

———. 2010. Highlighting E-Readers. *Inside Higher Ed*, February 23. Available at http://www.insidehighered.com/news/2010/02/23/ereaders.

Koninklijke Bibliotheek. 2009. e-Depot and Digital Preservation. Available at http://www.kb.nl/hrd/dd/index-en.html.

Krieger, Zvika. 2008. An Academic Building Boom Transforms the Persian Gulf. *Chronicle of Higher Education*, March 28. Available at http://chronicle.com/article/An-Academic-Building-Boom-T/20922/.

Kronenfeld, Michael. 2005. Trends in Academic Health Sciences Libraries and Their Emergence as the "Knowledge Nexus" for Their Academic Health Centers. *Journal of the Medical Library Association* 93(1): 32–9.

Kyrillidou, Martha, and Les Bland. 2009. *ARL Statistics 2007–08*. Washington, DC: Association of Research Libraries. Available at http://www.arl.org/stats/annualsurveys/arlstats/arlstats08.shtml.

Lardinois, Frederic. 2010. Amazon Now Sells 6 Kindle Books for Every 10 Physical Books When Both Editions Are Available. *ReadWriteWeb*, January 29. Available at http://www.readwriteweb.com/archives/amazon_now_sells_6_kindle_books_for_every_10_physi.php.

Lewis, David W. 2007. A Strategy for Academic Libraries in the First Quarter of the 21st Century. *College & Research Libraries* 68(5): 418–434.

Li, Chan. 2009. UC Libraries Springer eBooks Pilot. *California Digital Library*, February 2. Available at http://www.cdlib.org/inside/projects/springerebooks/.

Liu, Ziming. 2006. Print vs. Electronic Resources: A Study of User Perceptions, Preferences, and Use. *Information Processing & Management* 42(2): 583–592. doi:10.1016/j.ipm.2004.12.002.

Lloyd, Sara. 2009. Revisiting a Publishing Manifesto–What Does the Future Look Like for Publishers? *thedigitalist.net*, October 19. Available at http://thedigitalist.net/?p=714.

LOCKSS. 2008a. LOCKSS (Lots of Copies Keep Stuff Safe). Available at http://lockss.org/lockss/Home.

———. 2008b. "Springer Helps Launch CLOCKSS," news release, December 3. Available at http://www.lockss.org/lockss/Springer_helps_launch_CLOCKSS.

Lonsdale, Ray, and Chris Armstrong. 2008. Aggre-Culture: What Do E-book Aggregators Offer? *Library and Information Update* 7(4): 28–33.

Lown, Cory, and Hillary Davis. 2009. Are You Worth It? What Return on Investment Can and Can't Tell You About Your Library. *In the Library with the Lead Pipe*, April 1. Available at http://www.inthelibrarywiththeleadpipe.org/2009/are-you-worth-it-what-return-on-investment/.

Luther, Judy, Linda Bills, Amy McColl, Norm Medeiros, Amy Morrison, Eric Pumroy, and Peggy Seiden. 2003. *Library Buildings and the Building of a Collaborative Research Collection at the Tri-College Library Consortium*. Washington, DC: Council on Library and Information Resources. Available at http://www.clir.org/pubs/reports/pub115/contents.html.

Lynch, Beverly P., Catherine Murray-Rust, Susan E. Parker, Deborah Turner, Diane Parr Walker, Frances C. Wilkinson, and Julia Zimmerman. 2007. Attitudes of Presidents and Provosts on the University Library. *College & Research Libraries* 68(3): 213–227.

Lynch, Clifford. 2001. The Battle to Define the Future of the Book in the Digital World. *First Monday* 6(6). Available at http://www.firstmonday.org/issues/issue6_6/lynch/#l2.

Malpas, Constance, and Dennis Massie. 2009. Managing Shared Print Collections: From Evidence to Action. Presented at the Research Libraries Group Partnership Meeting, Boston, MA, June 1, 2009.

Mann, Thomas. 2001. The Importance of Books, Free Access, and Libraries as Places—and the Dangerous Inadequacy of the Information Science Paradigm. *The Journal of Academic Librarianship* 27(4): 268–281. doi:10.1016/S0099-1333(01)00214-2.

Manoff, Marlene. 2001. The Symbolic Meaning of Libraries in a Digital Age. *portal: Libraries and the Academy* 1(4): 371–381.

Marcum, Deanna. 2008. *Response to On the Record: Report of the Library of Congress Working Group on the Future of Bibliographic Control*. Washington, DC: Library of Congress. Available at http://www.loc.gov/bibliographic-future/news/LCWGResponse-Marcum-Final-061008.pdf.

Martell, Charles. 2007. The Elusive User: Changing Use Patterns in Academic Libraries 1995 to 2004. *College & Research Libraries* 68(5): 435–444.

MetaArchive Services Group. 2009. The MetaArchive Cooperative. Available at http://www.metaarchive.org/.

Mills, Keren. 2009. *M-Libraries: Information Use on the Move*. The Open University, University of Cambridge. Available at http://arcadiaproject.lib.cam.ac.uk/.

Montgomery, Carol, and J. A. L. Sparks. 2000. The Transition to an Electronic Journal Collection: Managing the Organizational Changes. *Serials Review* 26(3): 4–18.

Montgomery, Carol Hansen, and Linda S. Marion. 2002. Print to Electronic: Measuring the Operational and Economic Implications of an Electronic Journal Collection. *Learned Publishing* 15 (April): 129–136.

Nadler, Judith. 2008. Director's Message. The University of Chicago Library. Available at http://mansueto.lib.uchicago.edu/director.html.

National Endowment for the Humanities. 2009. Digging Into Data Challenge. Available at http://www.diggingintodata.org/.

Neilan, Catherine. 2010. Academic Publishers Seeing Strong Growth From E-book Sales. *The Bookseller*, March 10. Available at http://www.thebookseller.com/news/114429-academic-publishers-seeing-strong-growth-from-e-book-sales.html.

Nelson, Mark. 2008. E-Books in Higher Education: Nearing the End of the Era of Hype? *EDUCAUSE Review* 43(2). Available at http://connect.educause.edu/Library/EDUCAUSE+Review/EBooksinHigherEducationNe/46314?time=1205506595.

Newman, Michael. 2010. HighWire Press 2009 Librarian eBook Survey. High Wire Press. Available at http://highwire.stanford.edu/PR/HighWireEBookSurvey2010.pdf.

Nicholas, David, Ian Rowlands, David Clark, Paul Huntington, Hamid R. Jamali, and Candela Ollé. 2008. UK Scholarly E-book Usage: A Landmark Survey. *Aslib Proceedings* 60(4): 311–334. doi:10.1108/00012530810887962.

NYU Abu Dhabi Library. 2010. NYU Abu Dhabi Library. Available at http://nyuad.nyu.edu/academics/library.html.

NYU Press. 2009. NYU Press Leads Group Receiving Mellon Grant for UP Electronic Book Project. *From the Square*, June 26. Available at http://www.fromthesquare.org/?p=563.

OCLC. 2010. Cataloging and Metadata [OCLC–Products and Services]. *Cataloging and Metadata*. Available at http://www.oclc.org/us/en/services/cataloging/default.htm.

Ober, John. 2000. Library Services at California State University, Monterey Bay. In *Building Libraries for the 21st Century: The Shape of Information*, ed. Terry Web, 122–137. Jefferson, NC: McFarland.

Oder, Norman. 2009a. At BYU, Kindle Program on Hold, but University of Nebraska-Omaha's Program Going Strong. *Library Journal*, June 17. Available at http://www.libraryjournal.com/article/CA6666004.html.

———. 2009b. No, Abu Dhabi Is Not Paying To Digitize All of NYU's Library Holdings. *Library Journal*, October 20. Available at http://www.libraryjournal.com/article/CA6702879.html.

Oliver, Kathleen Burr. 2005. The Johns Hopkins Welch Medical Library as Base: Information Professionals Working in Library User Environments. In *Library as Place: Rethinking Roles, Rethinking Space*, 66–75. Washington, DC: Council on Library and Information Resources. Available at http://www.clir.org/pubs/reports/pub129/oliver.html.

Olsen, Stefanie. 2005. The College Library of Tomorrow. *CNet News*, August 3. Available at http://news.cnet.com/The-college-library-of-tomorrow/2100-1025_3-5817291.html.

OnDemandBooks. 2009. Our EBM Locations. Available at http://www.ondemandbooks.com/our_ebm_locations.htm.

Open Archives Initiative. Open Archives Initiative Protocol—Object Exchange and Reuse. Available at http://www.openarchives.org/ore/.

Orens, Anne. 2009. "Business Issues and Trends in the Digital Book Landscape: Players, Platforms and Devices." Presented at the NISO Webinar: E-Books: A Rapidly Evolving Marketplace, August 12, 2009. Available at http://www.niso.org/news/events/2009/ebooks09/.

Paulson, Kari. 2008. EBL ebooks can now be downloaded to the Sony Reader. *Ebook Library Blog*, October 29. Available at http://blog.eblib.com/?p=5.

Payne, Lizanne. 2007. *Library Storage Facilities and the Future of Print Collections in North America*. Dublin, OH: OCLC. Available at www.oclc.org/programs/publications/reports/2007-01.pdf.

Planets. 2007. Planets, Preservation and Long-term Access through Networked Service. Available at http://www.planets-project.eu/.

Portico. 2009. New Preservation Activities. *Portico*, August 14. Available at http://www.portico.org/news/preservation.html.

Prabha, Chandra. 2007. Shifting from Print to Electronic Journals in ARL University Libraries. *Serials Review* 33(1): 4–13. doi:10.1016/j.serrev.2006.12.001.

Price, Jason S., and John McDonald. 2008. To Supersede or Supplement: Profiling Aggregator E-book Collections vs. Our Print Collections. Paper presented at the 28th Annual Charleston Conference, Charleston, SC, November 6, 2008. Available at http://ccdl.libraries.claremont.edu/cdm4/item_viewer.php?CISOROOT=/lea&CISOPTR=161&CISOBOX=1.

———. 2009. Beguiled by Bananas: A Retrospective Study of the Usage and Breadth of Patron- vs. Librarian-acquired E-book Collections. Paper presented at the 29th Annual Charleston Conference, Charleston, SC, November 5, 2009. Available at http://ccdl.libraries.claremont.edu/cdm4/item_viewer.php?CISOROOT=/lea&CISOPTR=177#metajump.

Primary Research Group. 2008. *Library Use of E-books*. New York: Primary Research Group, Inc.

Princeton University. 2010. *The E-reader Pilot at Princeton (Summary Version)*. Available at http://www.princeton.edu/ereaderpilot/eReaderFinalReportShort.pdf.

Rajasekar, Arcot, Michael Wan, Reagan Moore, Wayne Schroeder, Sheau-Yen Chen, Lucas Gilbert, Chien-Yi Hou, et al. 2010. *iRODS Primer: Integrated Rule-Oriented Data System*. San Rafael, CA: Morgan & Claypool Publishers. Available at http://www.morganclaypool.com/doi/abs/10.2200/S00233ED1V01Y200912ICR012.

Rand, Deb. 2009. Results of Deb Rand's Survey. *NEEHSLs* and Beyond*, August 12. Available at http://newemergingmedicallibraries.wordpress.com/2009/08/12/results-of-deb-rands-survey/.

Reilly, Bernard F. 2003. *Developing Print Repositories: Models for Shared Preservation and Access*. Washington, DC: Council on Library and Information Resources. Available at http://www.clir.org/pubs/reports/pub117/contents.html.

Rich, Motoko. 2009a. The Vook: A Picture Book, but the Pictures Move. *ArtsBeat Blog, NYTimes.com*, September 30. Available at http://artsbeat.blogs.nytimes.com/2009/09/30/the-vook-a-picture-book-but-the-pictures-move/.

———. 2009b. Publishers Delay E-Book Releases. *The New York Times*, December 10, sec. Books. Available at http://www.nytimes.com/2009/12/10/books/10arts-PUBLISHERSDE_BRF.html.

Rothenberg, Jeff. 1998. *Avoiding Technological Quicksand*. Washington, DC: Council on Library and Information Resources. Available at http://www.clir.org/pubs/reports/rothenberg/introduction.html.

Safley, Ellen. 2006. Demand for E-books in an Academic Library. *Journal of Library Administration* 45(3): 445. doi:10.1300/J111v45n03_09.

Schonfeld, Erick. 2009. Kindle Notes and Highlights Now Accessible on the Web. *TechCrunch*, May 26. Available at http://www.techcrunch.com/2009/05/26/kindle-notes-and-highlights-now-accessible-on-the-web/.

Schonfeld, Roger. 2007. Getting from Here to There, Safely. *The Serials Librarian* 52(1–2): 183–189. doi:10.1300/J123v52n01_15.

Schonfeld, Roger, and Eileen Gifford Fenton. 2005. Digital Savings. *Library Journal* (March 1). Available at http://www.libraryjournal.com/article/CA504648.html.

Schonfeld, Roger, and Ross Housewright. 2008. *ITHAKA's 2006 Studies of Key Stakeholders in the Digital Transformation in Higher Education*. New York: ITHAKA. Available at http://www.ithaka.org/research/Ithakas%202006%20Studies%20of%20Key%20Stakeholders%20in%20the%20Digital%20Transformation%20in%20Higher%20Education.pdf.

———. 2010. *Faculty Survey 2009*. New York: ITHAKA S+R. Available at http://www.ithaka.org/ithaka-s-r/research/faculty-surveys-2000-2009/faculty-survey-2009.

———. 2009. *What to Withdraw: Print Collections Management in the Wake of Digitization*. New York: ITHAKA S+R. Available at http://www.ithaka.org/ithaka-s-r/research/what-to-withdraw.

Schottlaender, Brian E. C., Gary S. Lawrence, Cecily Johns, Claire Le Donne, and Laura Fosbender. 2004. *Collection Management Strategies in a Digital Environment*. University of California Libraries. Available at http://www.ucop.edu/cmi/finalreport/.

Scott, Jeff. 2009. Sony Using Libraries to Kill the Kindle. *Gather No Dust,* August 25. Available at http://gathernodust.blogspot.com/2009/08/sony-using-libraries-to-kill-kindle.html.

Shelburne, Wendy Allen. 2009. E-book Usage in an Academic Library: User Attitudes and Behaviors. *Library Collections, Acquisitions, and Technical Services* 33(2–3): 59–72. doi:10.1016/j.lcats.2009.04.002.

Siegler, M. G. 2009. Steve Jobs Returns to the Stage with Some Big Numbers to Share. *TechCrunch*, September 9. Available at http://www.techcrunch.com/2009/09/09/steve-jobs-returns-with-some-big-numbers-to-share/.

Slater, Robert. 2009. E-books or Print Books,"Big Deals" or Local Selections—What Gets More Use? *Library Collections, Acquisitions, and Technical Services* 33(1): 31–41.

Spiro, Lisa. 2009a. Interview with Jason Price and John McDonald. July 13.

———. 2009b. Interview with Raymond Siemens. July 13.

———. 2009c. Interview with Steve Stratton. July 16.

———. 2009d. Interview with Michael Furlough. July 24.

———. 2009e. Interview with Monica McCormick. August 10.

———. 2009f. Interview with Steve Shorb. August 18.

———. 2009g. Interview with Sue Polanka. August 19.

———. 2009h. Interview with Constance Malpas. August 21.

———. 2009i. Interview with Bruce Miller. August 25.

———. 2009j. Interview with Jacque Doyle. September 11.

———. 2009k. Interview with Roger Schonfeld. September 14.

———. 2009l. Interview with Diane Harley, Sarah Earl-Novell, and Sophia Krzys Acord. September 15.

———. 2009m. Interview with Michael Kronenfeld. September 17.

———. 2009n. Interview with Bill Robnett. September 23.

———. 2009o. Interview with Dee Magnoni. September 25.

———. 2009p. Interview with Saeed FakhriRavari. September 28.

Springer. 2010. eBooks Go to Print—MyCopy Launched. Available at http://www.springer.com/e-content/ebooks/my+copy?SGWID=0-165802-0-0-0.

Stone, Brad. 2009a. Amazon Erases Orwell Books From Kindle. *The New York Times*, July 18, sec. Technology/Companies. Available at http://www.nytimes.com/2009/07/18/technology/companies/18amazon.html.

———. 2009b. Best Buy and Verizon Jump into E-Reader Fray. *The New York Times*, September 23, sec. Technology /Internet. Available at http://www.nytimes.com/2009/09/23/technology/internet/23ebooks.html?r=2.

Sutter, John D. 2009. The Future of Libraries, With or Without Books. *CNN*, September 4. Available at http://www.cnn.com/2009/TECH/09/04/future.library.technology/index.html.

Tenopir, Carol. 2008. Ebooks Arrive. *Library Journal*, February 1. Available at http://www.libraryjournal.com/article/CA6523478.html?q=ebooks.

Tenopir, Carol, Donald W. King, Peter Boyce, Matt Grayson, Yan Zhang, and Mercy Ebuen. 2003. Patterns of Journal Use by Scientists through Three Evolutionary Phases. *D-Lib Magazine* 9(5). doi:10.1045/may2003-king. Available at http://www.dlib.org/dlib/may03/king/05king.html.

University of Chicago Library. 2008. Frequently Asked Questions. The Joe and Rika Mansueto Library. Available at http://mansueto.lib.uchicago.edu/faq.html.

UC Merced. 2009. About UC Merced. Available at http://www.ucmerced.edu/about_ucmerced/.

UC Merced Library. 2009a. Library Facts. Available at http://ucmercedlibrary.info/about-the-library/library-facts.html.

———. 2009b. The Story of the Library. Available at http://ucmercedlibrary.info/about-the-library/the-story-of-the-library.html.

———. 2009c. UC Merced Library Assessment Report (Spring). Available at http://ucmercedlibrary.info/library-assessment-information/.

U.S. Department of Health & Human Services. 2009. NIH Public Access Policy. Bethesda, MD: National Institutes of Health. Available at http://publicaccess.nih.gov/policy.htm.

Weiner, Sharon Gray. 2003. Resistance to Change in Libraries: Application of Communication Theories. *portal: Libraries and the Academy* 3(1): 69–78.

Welch Medical Library. 2009. Collections of the Welch Medical Library and Its Satellite Locations. Available at http://www.welch.jhu.edu/about/collections.html.

Wellcome Trust. 2008. Authors' Guide and FAQ, September 23. Available at http://www.wellcome.ac.uk/About-us/Policy/Policy-and-position-statements/WTD018855.htm.

On the Cost of Keeping a Book

*Paul N. Courant and Matthew "Buzzy" Nielsen**

1. Introduction

Among the critical functions of research and academic libraries is preservation—keeping both the scholarly record and much of the associated cultural record fit for use over time. Until recently, maintaining this record entailed managing primarily printed works, of ink on paper, bound in book form. The advent of electronic texts poses a novel and expensive set of preservation problems for academic libraries that have been addressed by many current and recent studies on the cost of digital preservation.[1] The topic of this report is the cost of storing and using print in old-fashioned codex form.[2]

We have two motivations for doing this work. The first is something of a straw man: as librarians and their funders become increasingly aware of the daunting technical and economic problems associated with digital preservation, there is often a certain wistfulness for

* We are grateful to Charles Clotfelter and a number of anonymous referees for helpful comments. We are especially grateful to Kathlin Smith for superb project management and editing.

[1] Several high-profile projects have addressed the financial implications of digital preservation: JISC and the British Library's Life Cycle Information for E-Literature (LIFE) project (McLeod, Wheatley, and Ayris 2006; Wheatley et al. 2007; Ayris et al. 2008); the University of California Libraries' Collection Management Initiative (Schottlaender et al. 2004); and CLIR's *The Nonsubscription Side of Periodicals* (Schonfeld et al. 2004). Others have studied digital preservation in light of the growing problem of where to store analog materials (Chrzastowski, 2003; Cooper, 2006; Schonfeld et al. 2004). Finally, many current projects on the subject are sponsored by governments on several continents, including the National Digital Information Infrastructure and Preservation Program from the U.S. Library of Congress and the Blue Ribbon Task Force on Sustainable Digital Preservation and Access.

[2] We use the term *preservation* often throughout this paper. By this we mean the long-term maintenance of materials for scholarly purposes and ensuring future access to the cultural record.

the good old days of print. After all, we have been solving the problem of print preservation for centuries. Just because we understand how to preserve print well does not mean, however, that we know how to do it without cost. It behooves us to calculate and understand the cost of keeping not only works that are born (or as yet unborn) digital but also print works that are currently stored by research libraries, are held in buildings that will deteriorate over time, and that will eventually have to be replaced. For most books in libraries today, including bound print journals, which we include as books throughout this study, we have already paid to keep the materials accessible for users today. But the next round of bills to be paid, both for new space and for replacement of existing facilities, is foreseeable and real. We undertake this study in part to show the continuing cost of holding print—old and new.

The second reason for undertaking this study is to help libraries evaluate collection and preservation strategies going forward. In any plausible configuration, academic and research libraries will be called upon to preserve and make available both print volumes and electronic records.[3] At the same time, as more works are made available digitally, libraries will increasingly have to choose between keeping a given work in digital or print form—acknowledging that either can be converted to the other at some cost. Libraries will face continuing choices in collection management, and making choices well will require understanding the cost of different modes of keeping materials accessible. Thus we are motivated to supplement the burgeoning literature on the cost of holding electronic records with a review and an addition to an older literature on the cost of keeping and using print books (hereafter referred to as *pbooks* when it is important to distinguish them from electronic books, or *ebooks*).

We take seriously an implicit commitment to maintaining, in usable form, the works that research libraries hold today and the works that they will continue to acquire in the service of scholarship. The commitment to preservation is not time limited—the international complex of research libraries has taken on the obligation of saving materials that others do not or cannot save. To complicate matters, where reliable electronic copies of works exist—and the number is increasing by tens of thousands a week—the argument for research libraries to share a good deal of both digital and print collections becomes stronger. The kind of collaboration needed to take advantage of the opportunities for shared collections is always expensive. Because the payoff to collaborative collection strategies depends, among other things, on the storage and preservation costs that can be avoided by employing such strategies, we need to have a fairly clear

[3] Even in a world where almost all use is digital, print can serve as a backup that is subject to a different profile of risk than electronic records. Thus it makes sense to keep some print copies as insurance against loss of electronic records. Additionally, even for works that are of quite ordinary quality and purpose (for example, university press monographs of the 1940s), the original print version may prove to have value as an artifact. Finally, original print copies often have significant historical value beyond the nominal content that is recorded in their pages. Libraries are in part museums of print, for many good reasons.

picture of the magnitude of those costs.[4]

In this report, we aggregate prior research and other information on pbook storage and analyze and synthesize these studies, supplemented by our observations and experiences at the University of Michigan. Our work draws heavily on earlier studies, particularly those by Cooper (1989, 1991, 2006); Lawrence, Connaway, and Brigham (2001); and Reilly (2003). We expand upon these studies by analyzing the data they present and drawing new connections about the relationship among facility type, storage policies, how books are used, and cost.

2. What's Involved in Storing a Book?

Upon examining the cost of storing a book, or several million books, it becomes immediately apparent that decisions about storage should be based on the anticipated use of the book. At one extreme—representing the approach most research libraries took until about 20 years ago—pbooks are placed on fixed shelving in facilities near their users. There is a good deal of space between shelves and shelving units, and the climate is controlled to make it comfortable for users to spend time in the stacks finding, retrieving, and replacing the books. In this scenario, the books are stored so as to be easily accessible and usable.[5] The real estate occupied by these books is usually near the center of campus and is therefore among the most desirable and valuable of locations.[6]

At the other extreme, pbooks can be stored in highly compact configurations, usually off-site and only accessible with lead times ranging from several hours to a day or two. This configuration is easier on the books and is cheaper in terms of land rent and construction cost per book. But access is also sacrificed. Browsing the off-site collection is generally impossible because the books are not shelved by subject, and although the labor required for storage is less than that in a central facility, the cost of accessing a particular book is generally much higher. In these configurations, books are stored in an environment that favors preservation, with substantially reduced convenience for the user.

The trade-off between storage cost and access implicit in these two extremes poses a number of issues as we attempt to assess the cost of storing pbooks. For example, direct comparisons of the costs need to be adjusted for ease of use. Additionally, libraries can move

[4] Collaborative collection projects are already in place for both print and digital repositories. Among these projects are HathiTrust (http://www.hathitrust.org) and the Research Collections and Preservation Consortium (http://recap1.princeton.edu/about/general.html).

[5] We have oversimplified, of course. Even 20 years ago, compact shelving was not uncommon. Moreover, there continues to be an important differential between open- and closed-stack facilities; the latter are more expensive to operate because they require more staff for circulation and more waiting time for users, but they are easier on the books. These differences matter, but for now we ignore them.

[6] This choice of location stems from the fact that in the print world the physical library was perforce at the center of scholarly activity. Almost everyone needed to use the library's works, which were available only by direct physical access. The cost of supporting scholarly work was minimized by placing the library's intellectual assets in the geographic center of the user population.

their collections across different facilities over time. A library could put books in the central, fixed-stack facility (which we term *open-stack*, even though in some libraries it is not open to all users) for a time, and then move them to more-distant storage facilities for the longer term. It could also make sense to place some new acquisitions directly into storage facilities, in cases where a library holds collections of record containing items that are not expected to be used frequently, and whose use, if any, is as likely in the future as in the present.

It is interesting that this trade-off has no parallel with ebooks. An electronic copy of a book, once securely stored on a server with appropriate redundant backup, can be browsed (although differently than pbooks), searched, and read pretty much anywhere, and pretty much instantly. For electronic works, there is no equivalent to compact distant storage, provided that the library has the rights to use the electronic works.

Because there are many different ways to store pbooks, our inquiry into the cost of storing a book will yield highly variable conclusions. How the book is stored, and how it is to be used, currently and in the future, will determine cost, and the cost differences that we discuss in this report can vary by as much as a factor of 12, depending on the assumptions made.

3. Space, Time, and Money

The term *life cycle* refers to a sequence of events or stages in maintaining a resource and making it accessible.[7] There is an extensive literature on the life cycle of library materials, which delineates a predictable course of uses, actions, and associated costs. Many authors, including Lawrence, Connaway, and Brigham (2001) and Shenton (2003), have advocated the use of life cycle analysis. The ongoing LIFE Project[2] uses a sophisticated implementation of the life cycle approach in assessing the costs of library materials (Wheatley, Ayris, Davies, McLeod, and Shenton 2007). Life cycle costs are organized by activities that vary over time, with some predictability. Using this approach, the total cost of a library resource can be decomposed into six parts: creation or purchase; acquisition by the library; ingest (i.e., processing upon receipt of the item); production of relevant metadata; storage costs; and cost of access or use.[8]

In this paper we focus on storage costs, but we will take note

[7] We could quibble with the term *life cycle* on the grounds that the standard life cycle in biology invariably includes death, whereas the life course for many library materials is meant to include permanent preservation, or as close to permanent as can be contemplated.

[8] The LIFE Project (Ayris et al. 2008) uses an equation and a set of symbols as follows:
$L_T = C + Aq + I_T + M_T + BP_T + CP_T + Ac_T$
L represents the total cost. This cost is composed of creation/purchase (C), acquisition (Aq), ingest (I), metadata (M), bit-steam preservation (BP, called "storage" in the first phase of the project), content preservation (CP, previously called "preservation"), and access (Ac). Ongoing costs are calculated over a time horizon, T. Because LIFE focuses on digital media, its cost categories reflect this focus, but the framework is easily adapted to print.

repeatedly of the fact that ease of access is determined in part by methods of storage, such that there is often a trade-off between storage costs and access costs. It is relatively cheap to store materials that are rarely used; conversely, it is generally quite expensive to use materials that are stored in high-density facilities far from users. Storage costs are often invisible in the annual budget because they may be subsumed in other budget categories such as building construction, maintenance, cleaning, climate control, or other areas that are often seen as part of library overhead. These costs would be much lower if the library did not store millions of books.

Time is crucially important to the use of life cycle modeling and storage costs. The total cost of storage at any given time depends on the costs incurred up to that point as well as on those that are committed in the future.[9] Many of the elements of life cycle cost, such as creation/purchase, acquisition, and even metadata, represent one-time, or at least irregularly occurring, costs. Metadata updates, for instance, may occur haphazardly or only during major database upgrades. The costs for storage and preservation of pbooks, as well as for access, depend chiefly on how long materials are to be kept, how expensive they are to circulate, and how frequently they will be used. In many cases, the right time period for this analysis will be indefinite—as close to infinite as the library can get.[10]

The length of time one expects to store a pbook greatly affects its ultimate cost, and the annual costs may increase or decrease depending on how well the book was cared for in its early years and on the quality of the medium on which it was printed. It is relatively easy to study how much libraries spend on electricity, buildings, and staff. But time is arguably the most significant variable librarians must consider in conserving pbooks. When research libraries purchase pbooks, in most cases they implicitly commit to maintain them in perpetuity. Whether they keep a book for only 10 years or for its entire life, the ongoing costs to maintain it may, and likely will, far exceed the volume's initial purchase price. Indeed, as Lawrence, Connaway, and Brigham (2001) estimate in a study similar to ours, the storage costs of a pbook over time may exceed the purchase price by about 50 percent.

Time is particularly important because as it passes, libraries' responsibilities grow. We mean this not in the sense that libraries gain new missions—although they assuredly do—but because the corpus of work that libraries are charged to keep and make accessible expands. Librarians are responsible not only for materials that their own generation deems worth preserving but also for everything that

[9] The LIFE Project considers the life cycle costs over a specified period of time (e.g., period 0 to T) and sums the preservation costs from each year. Assessment of the present value of costs allows one to make economically meaningful comparisons of costs incurred at different times. We will describe and employ present value later in this paper.

[10] For this reason, as we have discussed above, we believe that *life cycle* is not the correct terminology. More accurately, we have an essentially infinite lifeline for each item, with different actions required over the passage of time.

the preceding generations did.[11] To manage this increased volume of material, research libraries must (in some combination) secure more resources for storage, choose to discard an ever-increasing volume of material, or increase the efficiency of their storage. One mechanism that would improve efficiency would be to reduce duplication across libraries, but this issue is beyond the scope of this paper.[12]

The decisions we make early in an information object's life cycle influence both the future accessibility of the item and the overall life cycle cost of maintaining it. For instance, if a library chooses to leave a set of important newspapers in a hot, humid boiler room for 30 years, the damage done to those papers cannot be undone, and the cost of making these now-brittle newspapers accessible increases dramatically. To coin a term, we might call this the "Clementine Principle": if libraries do not properly care for their materials, whether electronic or print, from the beginning, those materials may be lost and gone forever. The loss is much more serious if there is no duplicate elsewhere. The cost associated with this loss is not reflected in the life cycle equations. It is the loss of value that would have been available had the material been kept fit for use. Avoidance of such losses—that is, maintenance of the scholarly and cultural record—is central to the mission of research libraries.[13]

Our work attempts to find the most efficient use of libraries' limited monetary resources for storage, making no assumptions about the value of the information in any particular volume. Economists call this "cost-effectiveness analysis," in which we hold output constant (we are holding a book's worth of information, indefinitely, at a specified level of accessibility) and compare the cost associated with different storage modes. This technique is complicated in the case at hand because it is difficult to hold the output constant. In particular, the trade-off between cost and accessibility is at the heart of decisions that libraries must make with respect to print storage. Implicitly, then, we are asking the reader of this essay to judge the value of delivery time and ease of browsing. Regardless of how that trade-off is resolved, we can determine cost-effectiveness. For any level of accessibility over any time path, a major source of cost will be the infrastructure that preserves the collection: buildings, climate-control systems, and technology. All must be replaced at some point, and these replacement costs are part of the total cost of providing continual access.

As we have already seen, the cost of storing a book depends on how it is stored and its use over time. At one extreme, libraries could opt to store books in densely packed, climate-controlled warehouses.

---

[11] Libraries should reassess continually what they are storing, but even if they decide certain information is no longer worth keeping, the vastly expanding amounts of information being produced, combined with the need to maintain the historical record, all but guarantee a growing commitment to preservation.

[12] For more discussion of the issue of duplication across libraries, see Schonfeld and Housewright 2009.

[13] We do not have good measures of the benefits generated by libraries and archives, though it is not for lack of trying. See, for instance, Griffiths and King 1994, Ozdemiroglu and Mourato 2001, Aabø 2005, and Americans for Libraries Council 2007.

In such a case, storage costs would be relatively small. However, such facilities reduce accessibility to patrons because the warehouses might be off-site or require staff mediation for checkout. At the other extreme, books could be stored in traditional main libraries, with standard shelves and climate controlled for users' comfort. Patrons could easily browse the shelves and consult materials of interest, but the suboptimal storage climate and potentially heavier use of materials mean that their future accessibility may be compromised and that future restorative costs may be higher. The actual average cost of pbook storage will generally fall between the costs generated by the open, main-stack model and the closed storage facility model because libraries often employ mixed strategies over time.

4. Dealing with Costs Incurred at Different Times—Discounting and Present Value

How much does it cost to keep a book for a century or more? Over this period, it is likely that the building that houses the book will be replaced two or more times; that the roof will be replaced even more often; and that the book will spend part of its life in accessible stacks, part in compact shelves, and, perhaps, part in high-density storage. Each of those systems will be constructed and installed at different times. And the buildings will be heated and cooled, requiring the use of fuel and electricity, the prices of which will change over time.

Economists compare expenditures undertaken at different times by using a technique called discounting to calculate the present value of all of the expenditures. The present value, in turn, is defined as the amount of money that we would need today to undertake the entire future set of activities that is contemplated at an assumed interest rate. In the case at hand, the present value of storage costs associated with a book includes the amount of money that we would have to spend today in order to persuade a reputable contractor to guarantee delivery of the requisite buildings, maintenance, and associated services in perpetuity. Perhaps surprisingly, the relevant amount is not infinite.

Suppose, for example, that the cost of storing a book for a year in today's prices is $3.00. Suppose that the interest rate on federal inflation-adjusted bonds is 3 percent. The present value of storing a book in perpetuity is $3.00 divided by 3 percent, or $100.[14] Why does it work? Because the $100 is just enough so that at the 3 percent interest rate, it will generate $3 per year. This works in the first year, the second year, and each succeeding year, into perpetuity. To generate $3.00 a year in perpetuity at an interest rate of 3 percent per year, one needs $100. At the end of the first year, the investment pays $3.00 and the principal amount of $100 is still intact. The concept is similar to an endowment, where an organization uses the interest while leaving the principal untouched.

Thus we say that the present value of $3.00 a year in perpetuity,

[14] See Gramlich 1990, 93-97, for an explanation of why this calculation yields the correct present value.

at a 3 percent discount rate, is $100.[15] In this scenario, we should put aside $100 to store a single book in perpetuity. We will use this money to pay for storage and upkeep of the book; whatever is not being used to pay current costs will be invested in bonds to generate income for future upkeep. Notice that calculations of this kind are very sensitive to the assumed discount rate. If we used 1 percent, the present value would be $300. If we used 10 percent, the present value would be $30. Our method in this paper is to calculate a present value for each element of storage cost.

Fortunately, inflation, which is difficult to forecast, is relatively easy to deal with in calculations of this kind. Interest rates generally exceed inflation rates, meaning that a dollar invested today will be able to purchase more than a dollar's worth of goods and services in the future, even after accounting for inflation.[16] For example, if prices are rising at 3 percent a year and the market (or "nominal") interest rate is 6 percent a year, a dollar that is saved for a year will buy the same goods it could buy today with three cents left to use for other things. Alternatively, if we anticipate buying goods a year from now that cost a dollar today, an investment of about $0.97 today is all that will be required.

In the example given here, the real rate of interest is 3 percent: the nominal rate of 6 percent less inflation of 3 percent. In the literature on benefit-cost analysis, it is common to assume a real rate of interest of 5 percent (Gramlich 1990, 93). Any positive real rate implies that current dollars are worth more than future ones. Discounting by a higher real interest rate would mean that today's dollars are worth relatively more—the future is discounted more heavily. A lower discount rate would have the opposite effect. To be conservative, we will calculate costs in this paper using the standard real discount rate of 5 percent, 3 percent (our base value), and 1 percent. Using the 1 percent rate results in future costs being higher in today's terms. The current economic situation tends to support the use of lower rates. Over the past 10 years, the CPI, the generally accepted measurement for annual price inflation, increased on average by 2.6 percent per year (United States Bureau of Labor Statistics 2009). The return on 10-year U.S. Treasury constant maturities averaged 5.2 percent annually over the same period (United States Board of Governors of the Federal Reserve System 2009). Three percent thus represents a fairly conservative—and fairly realistic—return on investment.[17]

---

[15] The term *discount rate* denotes the rate at which future sums of money can be made directly comparable to current dollars. In almost all cases, the discount rate will be the same as the interest rate. In this paper we preserve conventional economics usage and use the more general term discount rate. For more information on discount rates, see Gramlich 1990, 92-99.

[16] The current economic crisis has produced some short-term and short-lived counterexamples that are best ignored.

[17] The difference between the CPI and the return on 10-year Treasury maturities suggests that we should be discounting at 2.6 percent, not 3 percent. We choose 3 percent because, in the long run, it better reflects a conservative overall expected return from investments, and a lower rate would lead to an even higher estimate for storage costs than we calculate. We also show our calculations under the assumptions of 1 and 5 percent. In a paper similar to ours, Lawrence, Connaway, and Brigham

The use of real interest rates avoids the complication of trying to forecast the rate of inflation. If the assumed real interest rate is 3 percent, at an inflation rate of 6 percent the market rate of interest will be 9 percent. If inflation is 1 percent, the market rate of interest will be 4 percent. For our purposes, all that matters is the difference between the market rate and the inflation rate, namely, the real rate.[18]

Even though prices on average increase at the rate of inflation, the price of specific categories of production may rise faster or slower than the average. This will be important for our analysis of construction costs, which historically have increased at rates greater than general inflation. For construction, we will estimate inflation relative to prices in general. That is, if we anticipate that inflation in construction exceeds the growth in the CPI by two percentage points a year, which has been the norm for several decades, we can build that assumption into our calculations of the present value of storage costs, and continue to express the present value in today's dollars. All of our calculations will be expressed in terms of what money buys in 2009.

Our calculations, unlike those of Lawrence et al. (2001), Schonfeld et al. (2004), and others, assume that a given pbook will be stored in perpetuity.[19] This may seem odd, given the fragility of paper, but we would argue that perpetual storage best captures the mission of libraries. Except perhaps in the case of duplicates or ephemeral materials, research libraries generally intend to store their materials for as long as the institution exists, and they often spend money restoring, preserving, and, when necessary, duplicating deteriorating materials. To the extent that restoration and duplication are important, the calculations we make here underestimate the cost of preserving pbooks.

In sum, the idea of present value is essential for the kind of analysis we undertake here because of the very long time periods under consideration. When expenditures are undertaken at different times we can use present value to make each of them commensurable. Construction of 100,000 square feet undertaken in 20 years has a present value of the sum required today to pay for the construction then, in today's dollars. Put another way, how much would we have to invest today to cover the cost in 20 years?

(2001) use a discount rate of 7.5 percent. They use it because it represents "the long-term average discount rate delivered by state and municipal bonds" (p. 547). While they do not specifically state as such, this rate represents a "nominal" interest rate, i.e., one that does not factor out normal price inflation. As mentioned, using a high discount rate downplays future costs compared with present ones. A real discount rate of between 3 and 5 percent would more accurately represent the relative values of present and future costs. If normal inflation is subtracted from Lawrence et al.'s discount rate, their real rate would fall between 4 and 5 percent.

[18] Defining i as the market interest rate, π as the inflation rate, and r as the real rate, the formal relationship is that $(1+i) = (1+r)(1+\pi)$. For small values of r and π this is well approximated by $i = r+\pi$.

[19] The assumption of storage in perpetuity does not affect greatly our calculations of present value relative to storage for, say 100 years. At a real discount rate of 3 percent, a dollar spent 100 years from now has a present value of 5.2 cents. For a given sum of money spent annually, approximately 95 percent of the present value of perpetuity is accounted for in the first 100 years. Thus, our analysis would be little changed if we looked at storing a book for 100 years versus essentially for eternity.

5. The Costs of Pbook Storage

Our strategy is to estimate the present value of each element of pbook storage (for example, construction, energy, curation, maintenance) and to combine these values to approximate the total cost. Rather than try to develop one best estimate, we offer a range of estimates, reflecting varying assumptions about particular cost elements and the way in which books are used and stored (for example, in open main stacks, in closed storage facilities, or in a combination of the two). We also discount at a low rate of 3 percent, the more conventional rate of 5 percent, and the very low rate of 1 percent. With these varied assumptions, we wind up with a large range of estimated costs.

It is important to reiterate that storage cost alone (cost being primarily dependent on books/square foot) does not determine how best to store pbooks. Implicit in the range of storage choices is a range of functionality and operating costs: more books per square foot of library space requires more time, staff mediation, and transportation to get a book to a patron.

Following Gramlich (1990, 93-97), the arithmetic of computing present values for perpetual flows of resources is straightforward. To estimate the total present value, we add the following elements, as shown in table 1, on a per-volume basis: construction cost, maintenance cost, cleaning and janitorial services, electricity (including heating and cooling), staffing, and expected costs of circulation, recognizing that many volumes are unlikely to circulate at all.[20]

We calculate the present discounted value of each of these costs under three different storage models, with a slight variation in one case. In all cases we update past estimates to 2009 dollars.

- *Standard open-stack facility:* We estimate costs for a typical main library, with standard subject-organized shelving, assuming the industry standard of 10 books per square foot (Leighton and Weber 2000, 178). Our calculations are primarily based on Cooper (1989, 1991).
- *High-density storage facility:* These estimates—based primarily on a CLIR survey of such facilities (Reilly 2003)—represent costs for warehouse-style shelving buildings, likely located off campus or in a remote part of it. We assume 150 books per square foot (McLaren 2004, 20).
- *Hybrid model:* We estimate costs for a model that more closely matches what most libraries do: keep pbooks in a standard facility for a time before shifting them to a high-density facility. We

[20] We use the following formula to calculate the total net present value of storing pbooks:

$$\sum_{t=0}^{\infty}\left(\frac{\textit{Construction \$/vol.}}{1.01^{t}}+\frac{\textit{Maintenance \$/vol.}}{1.03^{t}}+\frac{\textit{Cleaning \$/vol.}}{1.03^{t}}+\frac{\textit{Electricity \$/vol.}}{1.03^{t}}+\frac{\textit{Staffing \$/vol.}}{1.03^{t}}+\frac{\textit{Prob. of circulation/vol.}*\textit{Circulation cost/vol.}}{1.03^{t}}\right)$$

In words, this formula states that the net present value of storing a book is the sum over an infinite number of years of the cost elements we identified divided by the discount rate raised to the power of the time (the year). Construction cost differs slightly by being divided by 1.01, as we estimate construction costs increase at 2 percent annually over general inflation. Details on the calculations within each cost element are noted below.

estimate two variants of this model. In one variant, we assume that the books stay in the standard facility for 10 years; in the other, the books stay for 20 years. This model includes the costs of building the high-density facility and transferring materials to it.

We estimate costs under what we consider to be the most likely scenario: we assume that construction prices increase at 2 percent annually over inflation, and we use a real discount rate of 3 percent, implying that the interest rate is 3 percent greater than the inflation rate. All values are in 2009 dollars.

Our base case results are shown in table 1. The units are dollars per book, and all figures, except the final row, are present values for perpetual storage. Details of how we calculate various cost categories are found in the next section, The Critical Elements of Storage Costs. Total costs under different assumptions are shown in table 2. These varying assumptions are discussed in the section entitled Costs under Different Assumptions.

Cost Element	*Shelving Model*			
	Open Stack	**High Density**	**Hybrid (10 years in open stack)**	**Hybrid (20 years in open stack)**
Construction	108.51	16.40	32.36	43.21
Maintenance	16.69	1.24	5.66	8.99
Cleaning	3.64	0.28	1.32	2.09
Electricity (heating and cooling)	2.39	0.20	1.03	1.53
Base staffing	6.08	1.20	2.42	3.36
Circulation	4.58	9.45	8.19	7.25
Total	**141.89**	**28.77**	**50.98**	**66.43**
Annual Average	**4.26**	**0.86**	**1.53**	**1.99**

Table 1: Our best storage cost estimates (in 2009 US$)

6. The Critical Elements of Storage Costs

Our estimates combine six major cost elements—construction, maintenance, cleaning, electricity, staffing, and circulation—though these variables are by no means comprehensive. In the following paragraphs we explain why we included these costs and how we estimated them. We use general estimates for typical open-stack and high-density facilities. Specific cost elements, such as construction, may differ by geographic region. Main campus libraries may be more expensive to heat in the northern reaches of the country, while storage facilities will be relatively expensive to cool in the south. Despite these variations, we believe these costs provide a good framework with which to understand how costs differ between facilities.

6.1 Construction

We use two primary sources to estimate construction costs for library facilities to house pbooks: Cooper (1989, 1991) for the construction costs of standard open-stack facilities and Reilly's (2003) CLIR report for storage facilities. These sources include shelving as a part of construction costs.

Unlike the other variables in our storage cost formula, for our base case we discount construction costs at 1 percent rather than 3 percent. Historically, construction prices have risen at rates much higher than general costs (for example, the CPI). Using the Fisher construction price index maintained by the U.S. Census Bureau (2009), we estimate that construction costs rise at a rate that is about 2 percent greater than general prices. Discounting assumes one will be able to get a basic real return on money if it is invested (in our case, 3 percent). However, with costs such as those of construction, which increase at a greater rate than general inflation, the potential net return on investing those monies is less roughly by a factor of the difference between the annual percentage increase and general inflation. Thus, if construction prices rise at about two percentage points a year more than inflation and the real discount rate is 3 percent, savings made for future construction costs will yield only 1 percent a year.

Many studies—including those of Schonfeld et al. and of the LIFE Project[2]—have done an excellent job of estimating costs for storing pbooks (or at least printed journals and other similar print materials). However, these studies generally don't address the replacement cost of facilities that are involved in keeping resources indefinitely. In our calculations, we assume that buildings must be replaced every 40 years, the estimated useful life of a building according to the American Hospital Association's *Estimated Useful Lives of Depreciable Hospital Assets* (2004). Hence, in year 40, 80, 120, and so forth, the cost of building a new building is incurred. If we lengthened this time in recognition of typical university practice, the results would not change greatly, although the present value of our estimated space costs would fall. For instance, if we assume that buildings are replaced every 60 years, the net present value of storing a pbook in an open-stack facility is $112.52 versus $24.35 for a high-density facility, compared with $141.89 and $28.77, respectively, under an assumption of replacement every 40 years. While the values are lower under the 60-year replacement model, the difference in costs between the two facility types is still large.

Finally, we vary the way we calculate space costs for the hybrid model. Because a book will not stay in an open-stack facility for the entire life of the building, we calculate an annual rent that is consistent with our assumptions about construction cost and building life. We use that estimated rent to calculate the cost to use the space in the open-stack facility for the first 10 or 20 years.[21] The result is $1.52 annual rent (in 2009 dollars) per book over the period. At year 10 or 20,

[21] We have specified periods of 10 and 20 years for illustrative purposes. In reality, libraries typically move an item off-site after its circulation drops below a certain threshold, a figure that likely differs depending on the subject area.

the cost of constructing the high-density facility is incurred. That cost then repeats every 40 years, as previously discussed.

The hybrid model incurs an additional cost during the years the books are transferred, namely, the cost to select and transfer the materials to the storage facility. We estimate these costs at $3.99 per volume, using figures from Cooper (1991, 417) updated to 2009 dollars.

6.2 Maintenance

Buildings must be replaced, and before they are replaced they must be maintained: bricks need to be resealed; heating systems fail; windows need to be replaced. Using figures gathered from the University of Michigan's Buhr Shelving Facility, we estimate a basic annual maintenance cost (including labor and materials) for library buildings. We derive these estimates by averaging a five-year cross-section of Buhr's maintenance costs, breaking down those costs by the cost per square foot, and then estimating a per-book value by factoring in the number of books per square foot for the various shelving types.

6.3 Cleaning

Buildings must be cleaned as well. We estimate cleaning costs using a method similar to the one we used to calculate maintenance. Taking five years' worth of cleaning data from the Buhr facility, we averaged it, estimated a cost per square foot, and calculated a cost per book based on the storage capacity of various shelving types.

6.4 Electricity

Electricity is a critical portion of the operating costs of any library. It runs the lights that allow users to see materials, powers the computers used to catalog them, and maintains the climate necessary to preserve pbooks.

We use a 1999 poll by the U.S. Energy Information Administration (EIA) to estimate energy consumption. EIA polled organizations on their energy consumption, dividing their findings across a number of categories, including buildings of different types and sizes, calculated per square foot. Storage facilities are generally kept at temperatures more hospitable to pbooks, while the climate of main libraries is maintained for human comfort, so it makes sense that there would be a difference in energy usage. In our calculations, we assume that open-stack facilities have the same energy footprint as an education-style building; warehouse-style buildings, by contrast, fit with storage facilities using compact or high-density shelving systems.

Using EIA's figures (p. 188) and the estimated number of volumes per square foot, we estimate the kilowatt-hour (kWh)/pbook/year at 0.91 for open-stack facilities and at 0.06 for warehouse-style facilities. Assuming that educational institutions pay the commercial rate for electricity (an average of $0.1028/kWh in 2008), we calculate an annual average kWh/pbook value and use that figure to estimate our overall electricity costs. The costs are varied in the hybrid model according to where a book is stored and for how long.

6.5 Staffing

The number of staff required to maintain a building is significant. After space-related costs such as maintenance and construction, we estimate that staffing is the next-largest cost associated with pbook storage. One problem with measuring such costs is that it is difficult to isolate the staffing costs that are strictly associated with storage, for example, people responsible for stacks maintenance or reshelving. To attempt to isolate these costs we again use Reilly (2003), who counted the number of staff (measured by full-time employees) for the facilities he studied. Using these figures, we estimate an approximate number of annual staff hours spent per book and, using an hourly rate inclusive of salary and benefits, estimate an annual cost.

We use $27 per hour as a standard hourly rate, including benefits. We calculated this value using the Association of Research Libraries' (ARL) 2007–2008 data tables (2009). These tables provide data on salary expenditures for professional, support, and student staff in research libraries nationwide. We calculated a single average hourly rate of $19.98, weighted based on the proportion each category of staff represents in overall library salary expenditures. Unfortunately, ARL does not include benefits in their figures. We use a benefits rate of 33 percent over the salary, which is typical at the University of Michigan, and round up slightly, resulting in an hourly rate of approximately $27. We then used this figure to estimate an annual staffing rate per volume.

This annual cost represents a base level of staffing necessary per book. For each type of facility, we then subtract from this base level of staffing the amount required for circulation (see fig. 1). Circulation is what economists call a marginal cost,[22] that is, it increases only as usage or circulation of the books increases. By subtracting the circulation costs from the overall base staffing as determined from Reilly's figures, we can estimate a fixed level of staff required per volume for each type of facility, independent of how much the typical item circulates.

Determining the storage costs for a main open-stack facility is difficult because so many other things are going on in such a building. Staffing in such facilities includes many people—reference librarians, system administrators, managers—who are not strictly associated with the storage, retrieval, and circulation of items. We therefore draw on Reilly (2003) and assume that the overall circulation and base staffing cost per pbook will be approximately the same for a standard facility as for a facility for storage-related purposes only. As the Reilly data regarding staffing show (see fig. 1), the relationship between staffing and current holdings is remarkably consistent.[23] The need for staff does not drop off as facilities become larger, as might be expected. What makes staffing costs different is

[22] *Marginal cost* is a standard term in economics referring to a cost that changes as the quantity of a good or service delivered changes. Specifically, it is the change in cost resulting when quantity changes by one unit. In our case, this change refers to the cost of storing one additional pbook.

[23] For the statistically inclined, the correlation is 0.86 and the r^2 is 0.74.

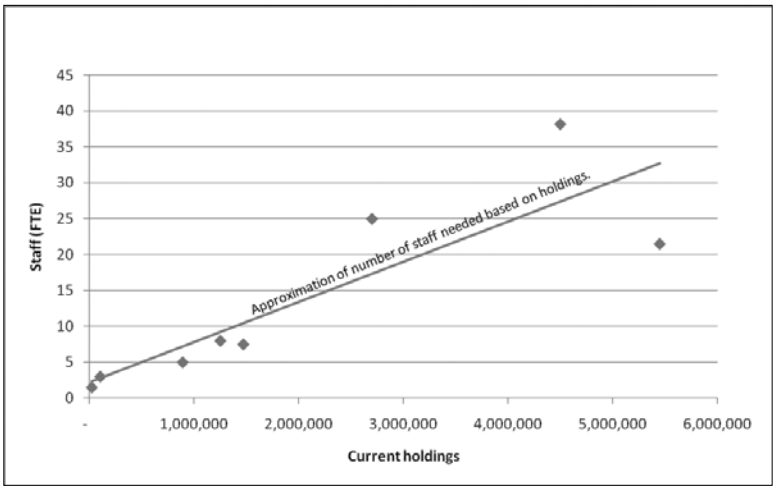

Fig. 1: Storage facility staffing versus current holdings.
Data Source: Reilly (2003)

circulation, which is more expensive per instance in a storage facility, but for which there are fewer instances per book. Open-stack facilities also have higher base levels of staffing, as they have higher use than do high-density storage facilities.

6.6 Circulation

Most of the cost elements we have mentioned dramatically favor high-density facilities. Circulation is the exception. Given the staff mediation and travel required, circulating an item from a high-density facility is much more expensive than from an open-stack facility. We estimate these circulation costs based on Cooper (1989), updated to 2009 dollars using the CPI and the producer price index (PPI) for gasoline costs (United States Bureau of Labor Statistics 2009). We assume a 25-mile round trip to deliver the materials to the patron.

To estimate an annual circulation cost, we take the updated Cooper figures and multiply them by estimated probabilities that an item will circulate from a specific facility. Payne (2007) reports that high-density facilities circulate about 1 to 2 percent of their collections annually. We use 2 percent. To estimate the circulation rate from open-stack facilities, we average the 2007–2008 ARL (2009) data on collection sizes and circulation, resulting in an annual circulation rate of approximately 13 percent. This figure may be slightly low for open-stack facilities, given that the ARL data include circulation from high-density facilities, but we believe it is a good approximation.

6.7 Other Factors

Readers may be thinking of other expenses, such as the cost of security systems and fire protection, replacement costs for climate-control systems, or insurance, that should be included in the cost of storing books. While these factors are important, we choose not to analyze them because their costs, when averaged per book, would be very low.

One element that may be worth investigating in future work, however, is the cost of replacing automated storage and retrieval systems (ASRs). These robotic order pickers are common in many newer high-density storage facilities (Boss 2002). While a building's useful life is estimated at 40 years, we suspect that the robotic order pickers in ASRs must be replaced before that. It is machinery, after all, and machines break down over time. Additional data on these replacement costs may change the cost differential between using an automated versus a human-mediated system.

7. Costs under Different Assumptions

The costs estimated earlier are those that we believe are most relevant to projecting future library storage costs, specifically the applicable discount rate, the relative inflation rate for construction costs, and the time that books spend in an open-stack facility before moving to high-density storage. Those three variables account for a fairly wide variation in cost estimates. We have already estimated costs based on whether books stay in an open-stack facility for 10 or 20 years. We estimate costs under slightly different assumptions in table 2. We discuss these differences below.

The discount rate dramatically affects costs because it has an impact on the weight placed on future costs. During the boom times of the late 1990s and the first several years of the current decade, discounting at 5 percent real interest would have been conservative; one could make far more than a 5 percent real return by putting money almost anywhere other than under a mattress. The recent financial climate is much less favorable for investment. Still, libraries are in the storage business for the very long term. What is a reasonable annual interest rate in the long run? We use 3 percent, which is quite

Assumptions	*Shelving Model*							
	Open Stack		**High Density**		**Hybrid (10 years in open stack)**		**Hybrid (20 years in open stack)**	
	Net present value	*Annual*	*Net present value*	*Annual*	*Net present value*	*Annual*	*Net present value*	*Annual*
Base	141.89	4.26	28.77	0.86	50.98	1.53	66.43	1.99
No construction increase	83.94	2.52	20.02	0.60	40.59	1.22	52.91	1.59
1% discount rate[24]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
1% discount rate and no construction increase	169.72	5.09	44.79	1.34	57.36	1.72	65.97	1.98
5% discount rate	73.12	3.66	15.37	0.77	43.99	2.20	61.76	3.09
5% discount rate and no construction increase	62.91	3.15	13.83	0.69	40.23	2.01	53.91	2.70

Table 2: Storage costs under different assumptions (in 2009 US$)

[24] It is impossible to calculate the net present value of storage cost under a 1 percent discount rate with construction costs increasing at 2 percent annually. This is because the construction costs increase by more than the amount of interest we could receive by investing the money, making construction costs essentially infinite.

conservative and which puts a fairly heavy weight on future construction costs. Five percent is more typical of analyses of this kind, so we make our calculations at 5 percent as well. For illustrative purposes, we also discount at 1 percent.

Our estimates are quite sensitive to assumptions about future construction costs. Implementing this assumption makes a larger difference on our cost than any other except the discount rate itself. Space is the single largest cost associated with storing books. Buildings are expensive, whether one pays rent to use them or pays for the property and construction outright. If we assume that space costs increase systematically relative to the overall price level, our estimate of the present value of storage will be higher than if we assume that the relative price of space is constant. Of course, construction costs do not always increase, as the financial events of late demonstrate, so we also provide estimates assuming no relative price increase for construction.

Most noticeable about the estimates for the hybrid model is the significant difference between keeping a book in open stacks for 10 or for 20 years, particularly at high discount rates. This difference arises from discounting and the fixed costs inherent to constructing facilities. Open-stack facilities are much more expensive to construct on a per-book basis. When books are kept in those facilities from the start, the highest costs are incurred in the early years, which contribute most to the present value.

The vast difference in costs between the hybrid model and both the high-density and open-stack facilities is also instructive. On the one hand, it tells us that we may be wasting money on items we do not expect to circulate often. Placing them in an open-stack facility is expensive, given the much higher per-volume cost of such facilities. If an item circulates infrequently, a high-density facility may save a great deal of money, even though circulating from that facility is 10 times as expensive as circulating from open stacks.

The degree of cost difference between the hybrid and "pure" models also suggests that, even if libraries put some high-circulation items in storage facilities, there still may be a significant cost saving. New storage facilities are expensive, but their cost pales in comparison with that of constructing a new main-campus, open-stack facility. Cooper (2006, 337) makes a similar argument with respect to bound journals. Substantial monetary savings are associated with using high-density storage facilities. Against these savings libraries must weigh the inconvenience and time costs imposed on users by slower retrieval and the inability to browse. In cases where electronic surrogates are available for pbooks held in high-density storage, the use of such surrogates can reduce these costs by providing alternative mechanisms for both browsing and retrieval of content.

8. Analysis

Tables 1 and 2 make clear that under any set of assumptions and any configuration of storage, the biggest costs derive from the

construction and operation of space. Dense storage is much cheaper than open-stack storage because it requires less construction, less electricity, less cleaning, less everything but staff per book. Circulating volumes from high-density storage is more expensive than circulating from low-density, so it is important to be careful about what sorts of materials are put into high-density storage.

High-density storage can be used to reduce costs, but a penalty is incurred in terms of functionality and circulation. There is a delay between ordering and obtaining a copy of a book; moreover, it is essentially impossible to browse a collection held in high-density storage. Thus our cost numbers are not really comparable (in the literal sense of being suited to make comparisons), and the cost advantage of high-density storage, while accurately portrayed for storage per se, does not take into account the operational disadvantages of high-density storage, which are often compounded by physical distance between the facility and the user.

Table 2 shows the effect of key assumptions on both the level of costs and the comparison of costs across different models of use. Increasing the discount rate reduces present values across the board, but has a smaller effect on annual cost and does not change the basic picture. The biggest effect would come from assuming that construction costs grow with inflation in general, rather than at a faster rate, as we have assumed. We see no basis for assuming such a favorable environment. Were it to materialize, all of our costs would fall substantially, because, as we have said, space costs are the principal driver.

The space costs that we have counted here do not include location rents. Including these would increase the dollar cost of all storage facilities, and would increase the cost of central campus facilities in the highest proportion. Central campus space is valuable for many purposes—classrooms, study and collaborative work space, arts production and display, administration, nearly every university function other than intercollegiate athletics and medical practice. If there were an active rental market within a university, the land upon which libraries tend to sit would be among the most expensive. Because we do not estimate land costs, we understate the true cost of holding books in open-stack facilities by a considerable amount. The economic advantages of high density and (as we discuss in the next section) electronic storage are even greater than the dollar estimates that we present here.

The cost advantages of off-campus high-density storage could be realized, at least to a substantial degree, through a complementary pair of strategies involving electronic storage and sharing of print collections. To the extent that digitized copies of print works are available to a university population for searching and browsing, it would be possible to restore much of the lost functionality that is inherent in high-density storage while retaining the cost advantages of such storage. (We are aware that the rights environment may limit that extent, and that the outcome of current lawsuits will bear upon it.) Users would search and browse electronically, eliminating or at

least reducing the need to make cursory uses of the physical book. Thus high-density storage would impose less cost in terms of functionality, and would likely be less costly to operate because there would be less circulation. The cost of running the electronic facility would have to be added, but as we will see, the net is likely to favor the kind of mixed-platform hybrid we suggest in this paragraph.[25]

Similarly, in an environment where there is widespread digital access, libraries could share their print storage, keeping only several copies nationally or regionally, rather than duplicating substantial swaths of their collections. Given the magnitude of the costs that we have discerned here, the savings from sharing of this kind could be substantial.

9. Comparison with Costs of Storing Electronic Books

A good deal of the current literature (Ayris et al. 2008; Beagrie, Chruszcz, and Lavoie 2008) shows that secure, long-term storage of digital objects is costly. Librarians bemoan the fact that these costs are often additional to print storage, in the sense that libraries will surely require the capacity for storage in both print and digital media.

As we briefly discussed earlier, however, for many titles libraries will have to choose between print and electronic copies. In many other cases, they will have no choice: vendors will provide one or the other. With respect to academic journals, the trend has clearly been toward electronic-only. Where there is choice, more and more libraries (Chrzastowski 2003; Johnson and Luther 2007) have switched exclusively to digital. The reason is often posed as usability. But considerable pressure and concurrent costs for storage have been removed, potentially reducing need for new facilities. Both functionality and storage costs are highly relevant to libraries' decisions about storage media.

Just as the question "What does it cost to store a pbook?" depends on how it is to be stored and used, so, too, does the question "What does it cost to store an ebook?" But the functionality of ebooks is much less dependent on storage than that of pbooks. To be sure, it is possible to put electronic resources into dark archives, but the darkness of the archive is not technologically determined; it is rather a matter of policy, usually as a result of copyright law, licensing agreements, or both.[26] When a library has rights to display the text of

[25] It is also possible that the ability to search and read electronically will increase demand for the physical resources. In this case, costs could rise because of the increased use unless libraries took offsetting actions.

[26] When the digital copy sits on a publisher's server and the publisher holds archival rights, the library's legal ability to assure permanent access is compromised. This set of problems is important and troubling (Jansen 2006; Stemper and Barribeau 2006) but in no way inherent to digital technologies. Several initiatives and organizations are working to ensure the future accessibility of digital content, including Lots of Copies Keep Stuff Safe (http://www.lockss.org), Portico (http://www.portico.org), and JSTOR (http://www.jstor.org).

a work, as in the case of public domain works, there is no electronic analog to off-site dense storage. On a server with redundant backup, the ability to search and read ebooks is essentially independent of the physical location of the server; users can access files from nearly anywhere with an Internet connection. Moreover, while pbooks deteriorate with use, the reliability of ebooks tends to improve with use. Even dark print archives—those that exist purely for backup—can be compromised in a variety of ways, intentional and accidental: for example, fire, flood, or poor stewardship. When something goes wrong with a collection that is being used, as with digital collections, the users can be relied upon to act as whistle-blowers. Since many more people are able to access files when they are provided digitally, there is an even greater chance that problems will be noticed.

The forms of electronic media relevant to the missions of academic libraries are growing and changing rapidly, and we have no way to predict how myriad elements of cost and functionality will play out. In this paper, we consider a relatively straightforward comparison—that of storage costs of a printed book versus the storage costs of page images and encoded text of the same book. We focus on relatively simple text and images, scanned or born digital, of the kinds that can be easily stored and retrieved in widely used formats, rather than on multimedia digital objects or databases. Many complexities regarding costs of ingest and development and reliable acquisition and production of metadata do not arise in this simple comparison.

Both the HathiTrust and the Internet Archive, among other entities, have a good deal of experience in storing electronic scans of print books.[27] HathiTrust provides rich access and reliable storage to ebooks at a fraction of our lowest estimates for providing compact off-site pbook storage. The predominant cost of print storage—space—is nearly absent for electronic storage, and the staff time devoted to electronic storage is less than that for storing and circulating print books. Moreover, and crucially, there is no reason to provide storage for ebooks that is difficult to access. Secure storage in the electronic case requires redundancy, which has no negative effect on access. Secure storage of print material makes access harder, rather than easier.

The HathiTrust provides a fully mirrored digital archive of millions of books, with tape backup, for less than $0.15 in fully loaded costs per book per year. Full color and a third site could increase the cost to as much as $0.40 cents per book per year.[28] Converting these

[27] See www.hathitrust.org for more information on the HathiTrust and www.archive.org for information on the Internet Archive. In both cases, explore the Web site and download and view public domain books to see the functionality provided by scanned texts.

[28] York 2009 provides documentation of the $0.15 annual cost for permanent storage per an OAIS Reference Model. (Downloaded from http://www.hathitrust.org/papers.) Per personal communication with John Wilkin, executive director of HathiTrust, and Paul Courant, founding and continuing member of the HathiTrust Executive Committee, these costs are fully loaded, including replacement of hardware and software and estimated costs of migration to new formats. $.40 per year is Wilkin's estimate of the upper bound on cost with an independent third site, again per personal communication with Wilkin.

costs into present value at the focal 3 percent discount rate that we have used in this paper would yield estimates of $5.00 and $13.10, respectively. Even $13.10 is less than half of the cost of high-density storage cost for pbooks shown in table 1, and is about a quarter of the cost for the most economical hybrid case. Moreover, it is likely that electronic storage costs will fall over time, which would reduce these estimates. Additionally, extensive use of ebooks for most purposes would enable libraries to use the most economical and secure methods for keeping reference pbook copies, and to share print storage as well. If everyone has a good electronic copy for use, it is not necessary for many libraries to hold print copies of the same works. A few instances of shared storage would do, as suggested in Schonfeld and Housewright (2009).

Storing and providing access to electronic material is indeed expensive and poses many problems, both technical and economic. And there is no doubt that complicated multimedia objects provide costly challenges to storage, some of which are not yet foreseen. But storage of scanned (or born-digital) books is much cheaper than equivalent storage of print materials. Where it is legally and functionally possible to make the move to electronic storage and use of the working copies of these kinds of materials, there is substantial economic gain.

The crucial differences in storage costs between electronic and print resources are found in expenditures for physical space and for access and delivery of works that are in high-density storage. In both of these domains electronic resources have enormous advantages. In table 3, we compare cost categories relative to the overall cost of print storage. Primarily because of the much smaller space required for it, electronic storage enjoys significant advantages. Even with very large collections of digitized works, the sheer amount of space required to store the servers on which those files reside will be dramatically smaller than that required for pbook storage. This factor alone will result in much lower costs, even if the cost per square foot of space is higher than for print storage.

Cost Element	Print	Electronic
Space	High	Much less
Cleaning	Low	Much less
Maintenance	Medium	Much less
Electricity/climate control	Low	Somewhat less
Staffing	Low	Somewhat less
Circulation/Access	Low	Much less

Table 3: Comparison of per-object cost of print versus electronic storage (relative to print cost)

10. Conclusion

Academic libraries will face many choices in the coming years as they continue to struggle with preserving and providing access to the cultural and scholarly records in an environment where the number and types of materials that they are expected to collect grow rapidly. As librarians grapple with these changes, it is important to recognize that the costs associated with a print-based world, often assumed to be small, are actually large.

Our analysis has been undertaken in terms of the cost of storage for a printed book. It makes the argument that the costs are high and that they sharply increase with the practical usability of a book. As we have seen with a number of journals, it is possible to substitute digitized print works for the original print (while keeping, as in the case of JSTOR, a set of original print copies). The savings in terms of space and the increase in functionality are parallel to the arguments made in this paper. If the cost of digitization is less than the difference in present value between print storage and digital storage, adding back in the cost of maintaining a shared print archive, there will be a net gain to the university sector of digitizing print collections and using the digitized versions for access. For most of our estimates of the cost of ebook and pbook storage, these conditions would hold. If another party, for example, Google or the Internet Archive, undertakes the digitization and provides the access, the argument becomes all the stronger.

Finally, we note that the argument in favor of moving toward digital versions of books and sharing both electronic and print collections is further enhanced when we recognize that university libraries tend to be located on prime real estate, and that there are uses of central campus stack space—for classrooms, study, offices, and enhanced library services, among others—that would be far more valuable than using that space to store materials most of which are used rarely, provided that access to the materials in aggregate could still be provided reliably.

Works Cited

Note: URLs are current as of May 15, 2010.

Aabø, S. 2005. The Value of Public Libraries: A Methodological Discussion and Empirical Study Applying the Contingent Valuation Method. Oslo: Unipub forlag.

Americans for Libraries Council. 2007. Worth Their Weight: An Assessment of the Evolving Field of Library Valuation. New York, NY: Americans for Libraries Council. Available at http://www.bibliotheksportal.de/fileadmin/0themen/Management/dokumente/WorthTheirWeight.pdf.

American Hospital Association. 2004. Estimated Useful Lives of Depreciable Hospital Assets. Chicago: Health Forum, Inc.

Association of Research Libraries. 2009. ARL Library Data Tables 2007–08. Washington, DC: Association of Research Libraries. Available at http://www.arl.org/bm~doc/08tables.xls.

Ayris, P., R. Davies, R. McLeod, R. Miao, H. Shenton, and P. Wheatley. 2008. *The LIFE² Final Project Report*. London, UK: LIFE Project. Available at http://eprints.ucl.ac.uk/11758/1/11758.pdf.

Beagrie, N., J. Chruszcz, and B. Lavoie. 2008. Keeping Research Data Safe: A Cost Model and Guidance for UK Universities. Available at http://www.jisc.ac.uk/media/documents/publications/keepingresearchdatasafe0408.pdf.

Boss, R. W. 2002. Automated Storage/Retrieval and Return/Sorting Systems. Chicago, IL: Public Library Association. Available at http://www.ala.org/ala/mgrps/divs/pla/plapublications/platechnotes/automatedrev.pdf.

Chrzastowski, T. E. 2003. Making the Transition from Print to Electronic Serial Collections: A New Model for Academic Chemistry Libraries? *Journal of the American Society for Information Science and Technology* 54(12): 1141–1148.

Cooper, M. D. 1989. A Cost Comparison of Alternative Book Storage Strategies. *Library Quarterly* 59(3): 239–260.

_____ . 1991. The Sensitivity of Book Storage Strategy Decisions to Alternative Cost Assumptions. *Library Quarterly* 61(4): 414–428.

_____ . 2006. The Costs of Providing Electronic Journal Access and Printed Copies of Journals to University Users. *Library Quarterly* 76(3): 323–351.

Gramlich, E. M. 1990. *A Guide to Benefit-cost Analysis*. Englewood Cliffs, NJ: Prentice-Hall.

Griffiths, J., and D. W. King. 1994. Libraries: The Undiscovered National Resource. In M. Feeney and M. Grieves, eds., *The Value and Impact of Information*, 79–116. East Grinstead, UK: Bowker-Saur.

Jansen, H. 2006. Permanent Access to Electronic Journals. *Information Services & Use* 26(2): 129–134.

Johnson, R. K., and J. Luther. 2007. The E-only Tipping Point for Journals: What's Ahead in the Print-to-Electronic Transition. Washington, DC: Association of Research Libraries. Available at http://www.arl.org/bm~doc/Electronic_Transition.pdf.

Lawrence, S. R., L. S. Connaway, and K. H. Brigham. 2001. Life Cycle Costs of Library Collections: Creation of Effective Performance and Cost Metrics for Library Resources. *College & Research Libraries* 62(6): 541–553.

Leighton, P. D., and David C. Weber. 2000. *Planning Academic and Research Library Buildings*, 3rd ed. Chicago: American Library Association.

McLaren, M. 2004. Current Library Collection Storage Models. *Kentucky Libraries* 68(1): 15–20.

McLeod, R., P. Wheatley, and P. Ayris. 2006. Lifecycle Information for E-literature: Full Report from the LIFE Project. London, UK: LIFE Project. Available at http://eprints.ucl.ac.uk/1854/1/LifeProjMaster.pdf.

Ozdemiroglu, E., and S. Mourato. 2001. Valuing Our Recorded Heritage. Paper presented at the Economic Valuation of Cultural Heritage Conference, London, February 2, 2001. Available at http://www.uni-hamburg.de/Wiss/FB/15/Sustainability/CCEMpaper.pdf.

Payne, L. 2007. Library Storage Facilities and the Future of Print Collections in North America. Dublin, OH: OCLC Online Computer Library Center, Inc. Available at http://www.oclc.org/programs/publications/reports/2007-01.pdf.

Reilly, B. F., Jr. 2003. Developing Print Repositories: Models for Shared Preservation and Access. Washington, DC: Council on Library and Information Resources. Available at http://www.clir.org/pubs/reports/pub117/pub117.pdf.

Schonfeld, R. C., and R. Housewright. 2009. What to Withdraw? Print Collections Management in the Wake of Digitization. New York, NY: ITHAKA. Available at http://www.ithaka.org/ithaka-s-r/research/what-to-withdraw.

Schonfeld, R. C., D. W. King, A. Okerson, and E. G. Fenton. 2004. The Nonsubscription Side of Periodicals: Changes in Library Operations and Costs between Print and Electronic Formats. Washington, DC: Council on Library and Information Resources. Available at http://www.clir.org/pubs/reports/pub127/pub127.pdf.

Schottlaender, B. E. C., G. S. Lawrence, C. Johns, C. Le Donne, and L. Fosbender. 2004. Collection Management Strategies in a Digital Environment. Oakland, CA: Regents of the University of California. Available at http://www.ucop.edu/cmi/finalreport/cmireportfinal.pdf.

Shenton, H. 2003. Life Cycle Collection Management. *LIBER Quarterly* 13(13/4): 254–272.

Stemper, J., and S. Barribeau. 2006. Perpetual Access to Electronic Journals: A Survey of One Academic Research Library's Licenses. *Library Resources & Technical Services* 50(2): 91–109.

United States Board of Governors of the Federal Reserve System. 2009. Market yield on U.S. Treasury securities at 10-year constant maturity, quoted on investment basis. Available at http://www.federalreserve.gov/releases/h15/data/Annual/H15_TCMNOM_Y10.txt.

United States Bureau of Labor Statistics. 2009. Consumer Price Index. Available at ftp://ftp.bls.gov/pub/special.requests/cpi/cpiai.txt.

_____ . 2009. Producer Price Index–Commodities. Gasoline. Available at http://www.bls.gov/ppi/.

United States Census Bureau. 2009. Price Deflator (Fisher) Index of New One-Family Houses Under Construction. Available at http://www.census.gov/const/price_deflator.pdf.

United States Energy Information Administration. 1999. *1999 Commercial Buildings Energy Consumption Survey: Detailed Tables*. Available at http://www.eia.doe.gov/emeu/cbecs/pdf/alltables.pdf.

Wheatley, P., P. Ayris, R. Davies, R. McLeod, and H. Shenton. 2007. The LIFE Model v.1.1. London, UK: LIFE Project. Available at http://eprints.ucl.ac.uk/4831/1/4831.pdf.

York, Jeremy. 2009. This Library Never Forgets: Preservation, Cooperation, and the Making of HathiTrust. *Digital Library Archiving* (May).

Ghostlier Demarcations: Large-Scale Text Digitization Projects and Their Utility for Contemporary Humanities Scholarship

Charles Henry and Kathlin Smith

1. Introduction

The title for this chapter is taken from the last line of Wallace Stevens' haunting poem "The Idea of Order at Key West," which also inspired the title for this volume. The poem explores, in part, the urgency of imposing a humanly recognizable order on a rough-edged world. It is a theme that resonates throughout the following pages, which report on a study investigating the utility to scholars of recent large-scale book-digitization projects.

It is not hard to intuit the potential for scholarly research of digitizing millions of volumes of books. Access to a sweeping array of our cultural heritage, in a form that can be queried, interpreted, and reconstituted as new knowledge, could transform scholarship. Yet to best support research, these digital databases must be organized and built to reflect the methodologies and intellectual strategies of contemporary scholarship. Until recently, discussions of large-scale projects have tended to focus on copyright, technical, and quality issues; few have included the perspectives of scholars on how they use these resources.

With funding from The Andrew W. Mellon Foundation and in partnership with Georgetown University, in 2007 the Council on Library and Information Resources (CLIR) initiated a project to assess the benefits and limitations for scholarship of text corpora made available through selected large-scale digitization programs. CLIR commissioned scholars in four disciplines to conduct these assessments in 2008 and early 2009. Their findings are summarized in section 3 of this chapter, and appear in full online.

In September 2009, CLIR convened a meeting of scholars to discuss the implications of the research findings and to identify features of these text databases that were considered well suited for advanced

research topics, and those functions that required improvement in the design and future development of mass-digitization projects. The day's discussions extended well beyond the research findings themselves. Participants noted that a new, complex, virtual scholarly environment had emerged during the three years since the project was launched. They underscored that many parts of this environment need to be coordinated and made coherent; otherwise, future scholarship may be impeded by poorly conceived digital architecture that does not facilitate productive searching, semantic nuance, or interoperable use across datasets—knowledge organization that does not align with the needs of intellectual production within higher education.

Two interrelated topics were of primary concern: (1) the need for current and future mass-digitization projects to be organized and designed to reflect the methodologies and strategies of humanistic scholarship and inquiry, requiring a stronger and more sustained voice of scholars in the development of such projects; and (2) how these projects, in turn, will influence the way research is conducted and how new knowledge is discovered and promulgated.

Sections 4 and 5 of this chapter provide an overview of the meeting discussion, and recommendations and next steps.

Participants, Meeting on Large-Scale Text Digitization Projects and Their Utility for Contemporary Humanities Scholarship
Washington, DC, Sept. 30, 2009

Jonathan Culler
Professor of English and Comparative Literature
Cornell University

Helen Cullyer
Associate Program Officer
Scholarly Communications Program
The Andrew W. Mellon Foundation

Anthony Grafton
Professor of History
Princeton University

Charles Henry
President
Council on Library and Information Resources

Rick Luce
Vice Provost and Director of Libraries
Emory University

Stephen Nichols
Professor of French and Humanities
Johns Hopkins University

James O'Donnell
Provost
Georgetown University

Kathlin Smith
Director of Communications
Council on Library and Information Resources

Eleonore Stump
Professor of Philosophy
St. Louis University

Will Thomas
Professor of Humanities
University of Nebraska-Lincoln

Nancy Vickers
Past President
Bryn Mawr College

Steven Wheatley
Vice President
American Council of Learned Societies

2. Study Context

Work being done in the digital humanities—a rapidly developing field of study, research, teaching, and invention concerned with the intersection of computing and the disciplines of the humanities—provides the context for this study. Centering on the investigation, analysis, synthesis, and presentation of knowledge using computational media, digital humanities studies how digital media affect the disciplines and how knowledge is constructed, maintained, processed, understood, and communicated. Digital humanists create archival collections, databases, and digitized objects. They use computational methods to analyze humanities materials in digital form and address scholarly questions about these sources. They often present their scholarship only in digital formats, making it accessible electronically and, more important, able to be transmuted, multiplied, revised, and reconstituted.

Digital humanities teaching involves methodological questions, narrative theories, computational programming, technical writing, group projects, and digital media productions. Students learn by immersing themselves in the information and producing a tool, model, project, or representation. Scholars working in digital humanities have developed archives for research and teaching, combining sources and materials from multiple institutional repositories around a common theme, historical question, or conceptual framework. Their work often depends on open linkages among objects and the recombination of digital objects into interpretive digital presentations. Leading examples of this work include Zotero, the Valley of the Shadow Project, the Walt Whitman Archive, the Spatial History

Project, and Virtual Pompeii, among others.

The challenge before scholars now is to make connections among and within huge sets of digitized data and to create new knowledge from them. As scholars seek to add interpretive and analytical value to large digitization projects, whether Google's or, for example, Cornell University's Making of America collections, they will need access to both metadata and data. If, for example, a scholar wanted to plot instances of Black railroad labor, whether enslaved or free, in U.S. history between 1830 and 1865, he or she would want to use Cornell's Official Records of the War of Rebellion to track military correspondence in the Civil War, and Google's digital books collection to find annual reports of companies that employed Black labor. That scholar might also wish to combine the data and metadata from these searches with several other data sets to create an interpretive spatial time line. Proprietary digitization efforts require scholars to enter the data individually from the screen, a practice comically out-of-step with the digital age.

These scholarly needs—accuracy of data, portability, multimedia recombination, and multivalent outcomes—inform this inquiry and its conclusions.

3. Summary of Commissioned Research

The research comprised investigations by scholars in four disciplines: linguistics (Melissa Baralt); Latin American literature (Patricia A. Soler); history (Alan Gevinson); and media history and cultural studies (Dawn Schmitz). They were asked to summarize key methodological considerations in conducting research in their disciplines, and to then assess Google Book Search (GBS), Microsoft Live Search Books (MLSB), and other relevant sources such as American Council of Learned Societies' (ACLS) Humanities E-Book (HEB), Internet Archive's (IA) text search, and the antecedent Project Gutenberg, with an eye to how useful the search and retrieval mechanism was for their work. The scholars' research reports, available online at http://www.clir.org/pubs/abstract/pub147abst.html, provide a high level of detail that is discipline-specific. A summary of their findings follows.

3.1 General Findings: Search and Retrieval

- **Book availability**. Scholars were able to find between half and three-quarters of the titles they sought among the titles available collectively from GBS, MLSB, IA, and HEB.[1] For books published after 1923, however, the availability of titles in full text drops dramatically.[2]

[1] The lion's share of digitized text is available through GBS. As of late 2009, GBS had scanned more than 10 million volumes. At the time of its closing in 2008, MLSB had scanned about 750,000 volumes, IA had made available 1.8 million volumes from the public domain (MLSB scans are now available via IA), and HEB had made 2,200 volumes available.

[2] Titles published before 1923 are in the public domain. Most titles published since 1923 are still under copyright.

They are available primarily in limited preview or snippets.

- **Scan quality**. Poor scan quality was a frequent problem. Issues ranged from blurred texts caused by movement of the page during scanning, to text obstructions such as fingertip smudges or sticky notes, to missing pages. Scanning quality was further affected by heavily marked-up books, which impeded optical character recognition (OCR), thus reducing text searchability.
- **Metadata**. The extent and quality of metadata—the basic descriptive information about a digitized work—varied widely among, and even within, the digitizing projects. Most titles carried basic, if limited, metadata. Metadata accompanying titles digitized by the IA tended to be the most complete. Some titles carried helpful chapter listings or a table of contents with hyperlinks, while other titles did not. Intermittent problems ranged from incorrect spellings of authors' names, to truncation of long titles, to missing information such as year of publication or editor's name, to incorrect page links. In some cases, metadata tags were wholly inaccurate.[3]

  Neither GBS nor MLSB offers MARC records, which include Library of Congress Subject Headings, along with other information. The links provided by GBS that are intended to facilitate searching or identify related books are sometimes of limited use for scholars, as Gevinson notes, "because listings of 'key terms' often exclude terms that are more 'key' than those included; links for 'key terms' often do not lead to all pages in which terms are discussed; and on occasion the links are totally useless."
- **OCR searching**. The digitization projects reviewed offer both page images, to represent a book's layout, and a text underlay that can be searched via OCR. Searchability is critical to scholars, but OCR is often unreliable. Misreadings occur as a result of poor scan quality or poor book condition, or because of OCR's difficulty in reading typographical variations such as footnotes, italicized words, capital letters, and hyphens.[4] Non-roman characters or phonetic symbols are a particular problem for OCR. The OCR on GBS did not recognize any Greek characters or Arabic letters, and electronic searching of specific phonemes is not possible, thus limiting the search utility for researchers of works in non-roman languages.
- **Modes of access: full, limited, or snippets**. Works published before 1923 are usually available for full-text viewing because they are in the public domain; those published later are generally not. For works in copyright, viewers may be allowed a limited view of a book's content (usually 10 or 20 percent of the total pages) or they may be allowed to see only snippets (that is, portions of

[3] In the period since CLIR commissioned the four investigations, considerable attention has focused on problems with metadata in GBS. See, for example, Geoffrey Nunberg, "Google's Book Search: A Disaster for Scholars," in *The Chronicle of Higher Education*, Aug. 31, 2009; and Paul Duguid, "Inheritance and Loss? A Brief Survey of Google Books," in *First Monday*, 12(8) [6 Aug. 2007].

[4] See http://googleblog.blogspot.com/2009/09/teaching-computers-to-read-google.html.

pages). Whereas limited-page views can be very useful in determining whether or not a book is relevant and worth the effort to obtain in full, snippets are far less helpful. Dawn Schmitz points out that "in many cases, [the snippet] viewing mode fails because the queried term is not even part of the snippet provided."

3.2 Discussion of Findings

In discussing the research findings, participants in the September 2009 meeting on large-scale text digitization projects focused first on the obvious trade-off between quality and quantity. Desktop access to an unprecedented body of text comes at the price of some accuracy. Until scholars have more confidence in the accuracy or completeness of texts made available through the largest digitization projects, many will consider the electronic surrogates as an important research aid, but not necessarily a replacement for the original texts.

A related concern was the persistence and propagation of poor-quality texts in the online environment. One participant noted the parallel in our experience with electronic catalog information: cataloging errors are numerous and have been replicated through MARC records, making them nearly impossible to correct. The result has been a noticeable deterioration in the accuracy of citation.

Participants noted the collection-level challenges that come with large-scale digitization projects. One such challenge is uneven collections. Large scanning projects use a quantitative, rather than qualitative, approach to selecting what becomes part of a digital collection because targeted selection is too expensive. This has the potential to create, as one participant put it, "an indiscriminate vacuum," and raises questions about the quality of results in interrogating large digital corpora.

Related to the problem of uneven collections is the question of what is missing from the digitized collections. Institutional collections do not overlap as much as people might think. Moreover, it is possible that many important books are missing from the scanned collections because they were checked out or otherwise unavailable on the day scanning was done. We need to identify the lacunae and determine whether there are patterns in what is missing. It would be interesting to identify the overlap between Google, Early English Books Online (EEBO), and other projects.

The commissioned research shows that we must be cautious about the degree of confidence we place in the corpora of texts created to date by large-scale digitizing projects as a resource for scholarship over time—whether because of the quality or completeness of the digital surrogates, their usefulness to some disciplines (for example, those using non-roman scripts), or the ability to repurpose the texts or develop tools that would enhance their usefulness.

Ideally, large-scale digitizing projects will adapt their processes and technologies to continually ameliorate the quality and usability of the product in which they are investing so heavily. In September 2009, Google announced that it had acquired reCAPTCHA, a

technology that improves the process that converts scanned images into plain text, and which it will use in its book and newspaper scanning.

4. Discussion of Broader Themes

The meeting participants' review of the research findings led to a discussion of broader concerns relating to the impact of large-scale digitizing projects. These concerns included trends in institutional support for libraries, funding for smaller digitization projects, and the future of print publishing. Each of these areas is discussed briefly below.

4.1 Impact on Library Budgets

Participants noted that university and college trustees and presidents frequently assume that large-scale digitizing projects will provide all the resources that scholars, teachers, and students require, and will free universities from the need to maintain libraries. At a time of severely restricted budgets, this assumption makes it much easier to reprogram funds for maintaining collections to other programs and services outside the library or unrelated to the humanities.

Decision-makers need to understand that:

- **GBS and MLSB are not replacements for books**. GBS and MLSB lead users to books, but are rarely a source of complete texts for works published after 1922. To use the book, it may be necessary to buy it or obtain it from the library (with many books out of print, the latter may be the only source). Digital access has been shown to increase demand for older books by pointing users to resources that could not be easily found otherwise. These books must remain accessible and in usable condition; this has implications for preservation budgets as well.[5]
- **Teaching, learning, and research rely on far more than monographs or print surrogates**. Users demand—and libraries provide—a broad range of materials and services.
  — *Undergraduate teaching requires access to varied resources*. Many humanities faculty members are beginning to tailor assignments around databases such as Proquest Historical Newspapers, American Periodicals Series Online, or 19th Century U.S. Newspapers, which the library makes available. Faculty members are also increasingly building the large public digital libraries, such as American Memory and National Archives, into their syllabi.
  — *Libraries are repositories for digital research*. Libraries are increasingly serving as repositories of important digital research and projects that cannot be maintained elsewhere. For example, the Virginia Center for Digital History at the University of

[5] This is not to say that each institution must keep a copy of every book that has been digitized from its collection. Libraries may consider keeping a "best copy" or copies in shared repositories. Courant and Nielson address the economics of book storage in this volume.

Virginia—home of the Valley of the Shadow and other notable projects—will be shut down in August 2010. The library will become the center for this activity.

- **Libraries are repositories for unique books**. A single title may exist in different versions at different libraries. As noted in the commissioned research, some scholars' work requires accessing more than one version of a given title. It is important that libraries preserve this variety, and that the digital availability of one particular version does not threaten the retention of other versions of the same title.[6]
- **Google is a commercial entity**. Google's key strength is its ability to search and index huge bodies of data. It is not in the business of data curation, but it is nonetheless a business. Its current business model makes a vast array of services available for free, but Google could charge for its services if it wanted to. The uncertainty of this future warrants caution for those who come to depend on Google for access to scholarly materials.

4.2 Impact on Smaller Scanning Efforts

A second concern is that large digitization projects may lead institutions to stop funding smaller scanning projects, in the belief that they duplicate what the large projects are doing. Many of these smaller projects, however, may be better designed than large-scale projects for use in humanities scholarship. For example, they may focus on content other than monographs, or they may be structured to allow users to mine and recombine text in ways that the current products of mass-digitization projects do not.

4.3 Impact on Books in General

Today, most people prefer to access resources online rather than in physical form, because electronic access is more convenient or cheaper. Teachers use electronic versions of primary resources because they are less expensive, or free, but these resources may not be of high quality. Participants expressed concern about the future of print publishing. If we bypass presses and peer review, how do we establish "best editions," as exist in the print world? If print publishing goes under, large parts of the world will be at a disadvantage because of infrastructure limitations, proprietary licensing restrictions, and even national policies that limit online access.

5. Research Recommendations and Next Steps

Participants made several suggestions for additional research and next steps. Their recommendations were guided by five goals: (1) to

[6] See Robert Darnton, *The Case for Books: Past, Present, and Future*. New York: PublicAffairs, 2009.

engage scholars more fully with the issues attending mass digitization; (2) to improve understanding of the burgeoning digital environment and its implications for teaching and research; (3) to gain the support and expertise of scholarly societies in these efforts; (4) to increase awareness among university administrators of the advantages and challenges of mass digitization; and (5) to build a more effective research program to monitor, assess, and influence the development, organization, and utility of mass-digitization projects from the perspective of the humanities scholar.

5.1 Research Recommendations

- **Conduct a baseline study of resources across a variety of disciplines**. The research that CLIR commissioned focused on text data sets as they relate to literary and cultural studies, linguistics, and history; other disciplines rely on different digital resources for research and teaching. The purpose of such a study would be to explore whether there is a set of common challenges with respect to large-scale digitizing efforts. Based on these findings, researchers could decide whether these interdisciplinary needs might be addressed by a short-term strategy as well as by a parallel longer-term effort.[7]
- **Repeat scholars' reviews of digitized sources to monitor improvement over time**. User interfaces for large digital collections will likely improve with time. It would be instructive to create a baseline from the current studies and to replicate them. It might be possible to automate some of the research that reviewers did manually to facilitate replication. For example, pattern-recognition software might be used to flag the frequency of text obstructions such as thumb images or sticky notes.
- **Find ways to use metadata that has been created for print books to support digital versions**.
- I**dentify lacunae**. Identify what is missing in the corpora of digitized works. Map the books digitized by Google against known collections in major repositories to determine the scope of Google's digitization. This may be facilitated by OCLC's decision to add records to WorldCat that represent digitized books from the Google Books Library Project and the HathiTrust Digital Library.[8]
- **Specify quality standards**. Conduct reviews, comparable to those commissioned by CLIR, of other electronic collections, such as those offered by Chadwyck Healey or JSTOR. Explore what

[7] According to *The Chronicle of Higher Education*, Google has created a grant program to support humanities text-mining research. The program will include the following disciplines: anthropology, archaeology, classics, history, linguistics, literature, philosophy, and sociology. See Marc Parry, "Google Starts Program for Studies of its Digitized Books," *The Chronicle of Higher Education*, March 31, 2010. Available at http://chronicle.com/article/Google-Starts-Grant-Program/64891/. See also "Google Digital Humanities Research Awards: 4/15 deadline, $50K Awards," available at http://uodigschol.wordpress.com/2010/03/16/google-digital-humanities-research-awards-415-deadline/.

[8] Information on HathiTrust is available at http://www.hathitrust.org.

standards of quality control, access policy, audit frequency, and terms of sustainability are necessary to convince scholars that a digital object is an adequate, acceptable, and trusted surrogate for a printed book or article. These factors may vary among disciplines.

5.2 Next Steps

Participants identified three areas for continued work: (1) integrate the voice of scholars in the planning, development, and evolution of all major digital projects that support humanistic teaching and research; (2) encourage greater international cooperation on issues of large-scale digitization; and (3) support a more concerted effort to develop the digital architecture necessary to facilitate multivariate research, portability of data, and accuracy of information around the world, creating a digital environment that supports new methodologies, new questions, and new discovery.

Within this context, the participants made four recommendations:

1. Convene a task force to direct ongoing research, engage scholarly and professional societies, and communicate with campus leadership on the issues. The task force would meet regularly to commission reports and projects and to monitor and assess the advantages, risks, and implications of existing and emerging large-scale digitization projects. The American Council of Learned Societies was suggested as an ideal partner for the task force.
2. Create a centralized body, other than search engines, to provide guidance in the way that printed readers' guides did. A clearinghouse on the Web—a Reader's Guide to Digital Scholarship in the Humanities—could become indispensable for scholarship, and thus could have a greater chance of long-term sustainability than search engines. Such a guide might include reviews in selected academic fields—or blogs—to which people could subscribe, or an RSS feed to Zotero in particular fields. Professional societies could take on this role, or they could broker access to resources that are usually available only to research organizations. The task force might include a subcommittee of technical experts who would explore issues pertaining to digital architecture and knowledge organization over very large, independent data sets.
3. Engage humanities scholars internationally. Institutions in many other nations are digitizing on a large scale—the Koninklijke Bibliotheek in the Netherlands, the British Library, the National Library of China, the National Library of Germany, JISC, and Bibliotheca Alexandrina are just a few. The products of these digitization efforts may meet the standards and needs of scholars more satisfactorily than the products of U.S.-based projects. These national and continent-wide programs and projects need to be identified. Their standards and goals must then be assessed and, ideally, integrated into U.S. efforts.

4. Use the task force to engage in conversation and establish working relationships with developers of mass-digitization projects. An important purpose of this engagement would be to articulate the implications of such projects for research and teaching in the humanities, using this report as a basis for that dialogue. Large-scale corporate and private developers would gain insight into how their products are used and could tailor them more astutely to an academic market.

Conclusion

Roger C. Schonfeld

We are at a moment of great uncertainty as to the extent of the digital transition for scholarly resources. The three essays in this volume add significantly to what we know about the relationship between print and electronic versions of scholarly materials and the impact that this relationship might have on scholars and libraries. The essays provide a scholarly perspective on the transition from print to electronic format and analyze the managerial implications of such a transition. Taken together, they raise two related themes. The first is the tension that research libraries face between fulfilling their time-honored role as custodians of scholarship and enabling a digital environment for scholars. The second theme is the growing potential for systemwide analysis and response to help mitigate this tension.

A Formal Migration?

For many academic libraries, reducing the resources expended on processing and storing print presents both an opportunity and a challenge. The opportunity resides in the ability to allocate resources to new roles and services whose value to library stakeholders is increasing. The challenge comes from the perception, sometimes held by academic-resource allocators, that print no longer has any role at all. Libraries today strive to strike a balance between reducing print expenditures and ensuring that support for print is maintained where it is important. There are few models to follow, and although our experience in moving from print to electronic journals is instructive, we should not assume that it will be replicated with books.

In the case of scholarly journals, libraries stand to greatly reduce expenditures for print formats. It is not just that electronic storage costs will be lower than print storage costs, as Courant and Nielson

amply demonstrate, but that digital collections will be organized differently than print collections. A publisher or vendor typically provides access to e-journals from a central platform rather than an individual library, with this scale effect compounding the format cost differential. For digitized journal backfiles, the library community has come to recognize the need to retain print preservation copies on a shared basis, but not necessarily at individual libraries. For both print and electronic versions, storage and provision of journal collections are moving to the systemwide level. Libraries and publishers have collaborated to create a robust digital-preservation infrastructure alongside a program for audit and certification of this infrastructure. This should make scholars and libraries confident that born-digital materials will be assured of the same long-term preservation and accessibility that libraries have traditionally provided for print journals. As libraries increasingly substitute print journal acquisitions with licensed e-journals, publishers are carefully considering whether and when to wind down their print publishing programs. For scholarly journals in most fields, the dual-format period appears to be waning.

If the same pattern could be assumed for monographs, then the need for research libraries to maintain tangible collections of general materials would be called into question. But as Henry and Spiro rightly conclude, books possess a completely different set of affordances from the perspective of format migration. The authors give many reasons why newly published scholarly monographs in electronic format are not yet (and indeed may never be), a complete substitute for print—reasons that include the challenges associated with reading at length in electronic format, immature pricing and dissemination models, and an inadequate rights environment, among others. At the heart of these issues seems to be whether academia will ultimately find itself a small part of an e-book ecosystem defined by trade publishers, commercial Web search engines, and consumer retail opportunities, or in the position of being able to define that ecosystem, as it did for scholarly journals.

The scholarly publisher-driven model that has turned journal publishing electronic may not obtain for books. Amazon, Apple, and Google are responding in the first instance to the needs of trade, rather than scholarly, book publishers, and audiences. Moreover, while a relatively small number of large-scale online platforms have emerged for journals, the platform landscape for books is unresolved, and particular challenges may exist given the fragmented university press environment for scholarly monographs.

An equally complex set of issues obtains for the digitization of older print books. Google's initial announcement of its book-digitization project cited search as its primary motivation, and various elements of the project were apparently optimized toward this objective. There was no sense that this was a project with a primarily scholarly audience, or one for which format substitutability was a key objective. Still, public domain materials have long been readable through the Google Books interface, and the copyright lawsuit

settlement, if approved, would see Google selling access to many in-copyright books as well.

Even assuming that perfect scanning fidelity could be linked to outstanding metadata, Henry and Spiro suggest that it is far from clear whether digitized print books would offer a viable substitute for print originals. One key question is how much reading of print books is cover-to-cover and how much is targeted through entry points such as a table of contents or an index: that ratio would suggest the extent to which screen reading and printouts could be as valuable for books as they have been for journals. For cover-to-cover reading of long-form works, the burgeoning field of reader and tablet devices, if linked to collections of scholarly books, might draw still more users away from print versions. When it comes to digital versions of scholarly books, these key issues remain unresolved, and the prospects for a true format migration therefore remain uncertain.

Whether or not widespread format substitution is ever realized for scholarly books, electronic versions have already proven to be valuable complements to print. Search and discovery are being transformed by Amazon's and Google's full-text programs. More advanced forms of scholarship, reliant on technologies such as text mining and other computational analyses, are emerging across a variety of fields. Henry and Smith make clear that the quality of digitization is critical for these new uses, and that scholars may be most frustrated not at the failure for digital to substitute for print but rather at its failure to yield the complementary and innovative opportunities.

The appeal of e-books will continue to grow as reader devices become more pervasive and useful; as new business models and services are developed to support e-book discovery, access, and distribution; and as mass digitization becomes better suited to scholarly needs. But today, academia is not prepared for a wholesale transition away from print format for monographs. It is likely that a dual-format environment will obtain for books for the foreseeable future, forcing libraries to bear the costs of licensing and maintaining access to electronic versions as well as the costs of print. Because it will raise, rather than lower, overall costs for books, a dual-format environment will be a real challenge to libraries seeking to reduce collections costs and redirect resources to new roles and services.

For this reason, efforts to find greater flexibility in the management of monograph collections are worthy of experimentation. As interlibrary loan networks have grown, so has the need to find the right balance between making a local acquisition and relying on peer libraries for access to that item. In addition, efforts to share collections through collective purchasing and identifying unnecessary overlap in monograph collections should continue.

Building Digitized-Book Collections

The project management, technology, coordination, and funding that Google and other third parties have brought to bear have been

the most important factors in advancing digitization initiatives. The prospect of mounting a book-digitization project at the scale of Google's never seemed within the capabilities of research libraries until an outside partner with seemingly limitless resources emerged. But this outside partner has often seemed more focused on its commercial prospects than on meeting scholarly needs. Academia must meet its obligation to serve the needs of scholars within the context of a highly resource-constrained environment.

The pending settlement agreement between author and publisher plaintiffs and Google could be approved (or denied) any day. If approved, its impact would be transformational. Google's collection of digitized backlist books would achieve, or exceed, the scale of the largest journal-digitization initiatives. The corpus would be licensed to both academic and other types of libraries.

Because scholarly needs can often be more readily addressed by organizations and initiatives that are focused on academia, the creation of the HathiTrust and its relationship with Google is especially promising. In essence, Hathi will bring an enormous collection of digitized books, including many of those scanned through Google partnerships, under the control of an enterprise principally driven by a scholarly mission. Hathi's digital-preservation mission is vital, and its partnership model allows participating libraries to achieve scale and quality far greater than that which any of them could do on their own.

Opportunities associated with bringing the content in question under the control of academia may extend beyond preserving the digitized files themselves. The Hathi partnership might provide another forum for the reimagining of the preservation role for general collections in a digitized environment, where a key question is whether and how preservation responsibilities shift from a local to a systemwide level. For example, how much priority should libraries assign to the preservation of print general collections once they have been digitized, when they might be less frequently used or serve less of a preservation function? Can preservation resources be used to correct errors in digitized general collections (such as those preserved digitally in Hathi) to ensure that their quality is maximized? The preservation community has only begun to grapple with these questions.

The value of these many preservation and collection-management activities will vary depending on whether and how reading and other uses migrate to the digital format. It will be necessary to connect that value to the resources allocated to the various activities. For example, if book reading stands any significant hope of moving toward electronic interfaces broadly, Hathi and its partners could consider whether it would ultimately be more cost-effective for the library community to maintain a dual-format environment for books or to upgrade the digital versions to enable a transition. While it may be premature to contemplate such calculations seriously, the library community can gather data and begin to equip itself to conduct such analyses.

In Sum

This volume has established that the future of collections and the potential for format transitions will be influenced by a combination of scholarly needs and library and publisher management challenges. The order that we all seek in this unsettled environment is not, however, simply its own pursuit. Scholarly needs are shifting, the information ecosystem is changing rapidly, and libraries are finding that many of their traditional roles are no longer as highly valued as they once were. As libraries bring order to their collections by taking advantage of systemwide opportunities that are appropriate to their stakeholders, they will be able to bring additional innovation and agility in addressing new needs.

Epilogue

Charles Henry

In 2009, The Andrew W. Mellon Foundation awarded CLIR a grant to study the feasibility of inaugurating a "cloud library," defined as a library that is provided digital surrogates of copies held by other institutions and organizations. The term was taken from the concept of cloud computing, whereby an institution elects to have services, business applications, and resources supplied by another party through the Internet rather than by a locally owned and operated client server. The study explored the feasibility of such an arrangement with the aim of deaccessioning less-used titles in the client's holdings. It also examined the potential for relying on a holding in a nearby print book repository as backup.

The high costs of maintaining redundant print collections provided the impetus for the study. As Paul Courant and Matthew Nielson have shown in this volume, the cost of purchasing the materials is only a fraction of the investment needed to maintain the printed record over time. Construction costs, environmental upgrades, maintenance, and expansion of local storage space contribute to the considerable overhead of keeping the human record.

Significant technological advancements over the past several years have made new models of collection preservation and access feasible. Large-scale digitization of books provides an opportunity to create astonishingly vast repositories that duplicate large portions of research libraries' holdings. These projects also can offer more sophisticated means of searching and reconstituting information than does physical browsing. High-speed networks for rapid delivery, increased security, and analytic tools that facilitate new methods of inquiry and intellectual strategies augur for a more efficient and cost-effective model of custodianship for scholarship and research.

Initial activities associated with developing a prototype of a cloud library include conducting research to identify the policies,

procedures, logistics, and infrastructure necessary for the potential elimination of monographic printed titles duplicated electronically and in another accessible repository, at what could be a considerable savings over time. In the Mellon-funded study, New York University (NYU) served as the client library, the digital repository was the HathiTrust, and the off-site collaborative storage and retrieval project ReCAP was the shared print repository. The initial findings of the study show that relying on ReCAP and the HathiTrust is not tenable as a substitute for less-used titles at NYU. The overlap of titles in ReCAP was too meager, and accessible digital titles (i.e., those that could be accessed and read by NYU library users) currently available in the HathiTrust that overlapped NYU holdings were too few to serve as a viable alternative.

Nonetheless, the extensive data analysis performed in conjunction with the study identified some important issues that suggest next steps. Projections for the number of titles in the HathiTrust—approximately 14 million by 2012—include about 25 percent of the titles available in the public domain.[1] At that scale, the HathiTrust could serve as a surrogate research library of public domain titles for many, if not most, of the institutions in the United States. In addition, the study made it clear that ReCAP and similar repositories, unlike the HathiTrust, do not resemble research libraries. The former are ad hoc and eclectic in nature, offering little in the way of collection development strategy. Their holdings are also difficult to assess, since even repositories that are shared across institutions have catalogs that are subsets of the owning institution's collections and are not shared or interoperable with other electronic records. For example, a reader has to go to the Columbia University catalog for ReCAP titles, and then go through Princeton's catalog to determine what additional books might be at the facility that Princeton owns.

This study identified the need for two ambitious, large-scale projects. The first project is a rationalization of major U.S. printed book repositories. If selected off-site storage facilities could be identified and their electronic records federated into what would appear as a central database, then the existing off-site collections could be rationalized and reconfigured as a national system and redundant copies above an accepted base of duplications could be culled. As a national system, the repositories could also be logically developed and grown in the years ahead.

The second project is the formation of a large coalition of research libraries that would agree to use the HathiTrust as a digital surrogate for selected collections. These libraries would digitize, where necessary, and render to HathiTrust copies of books that could be subsequently shared by the consortium, while making possible a large-scale, coalition-wide deaccessioning of print books. The two projects are interrelated and would, taken together, create a trusted,

[1] To date, HathiTrust contains more than six million digitized volumes. A significant portion of the holdings contributed by partner libraries were digitized through Google Book Search; the projected growth in holdings is based largely on the additional scanning that Google expects to do in the next two years.

sustainable regime that balances digital and print resources.

The longer-term solution to the high cost of maintaining redundant print copies of books can be addressed only through a coordinated review of policy and practice at some of at the largest shared print repositories with an aim to address these national objectives. Our future, it appears, depends upon our ability to reconceptualize the traditional model of competing, stand-alone institutions into a coherent system that not only preserves the identity and independence of universities and colleges but also brings together many of the functions and support services that undergird scholarship and teaching in ways that are more effective, efficient, and elegant. As the essays in this volume have articulated, planning for and building this new digital commons is both an extraordinary opportunity and a complex challenge.