In Memoriam

Peggy Johnson with assistance from Carla Dewey Urban

Edward Swanson, *LRTS* Book Review Editor, died December 10, 2010, after a brief illness.

Edward's long and significant career (after earning his master's degree in library science from the University of Minnesota) began at Macalester College, his alma mater, with duties including cataloging and organizing the college archives. In 1968, he moved to the Minnesota Historical Society where he led the Newspaper, Processing, and Technical Services departments, culminating in the position of coordinator of Library Cataloging and principal cataloger. During this time, he played a vital role as a Minnesota AACR2 Trainer, helping librarians throughout the state learn and understand the new cataloging rules. He not only provided in-person training, but authored and edited numerous manuals and other documentation to support cataloging. Edward prepared curriculum and conducted training for the Minnesota Opportunities for Technical Services Excellence (MOTSE) continuing education program, strengthening the cataloging knowledge of librarians and paraprofessionals throughout the state. He also served as a long-time Name Authority Cooperative Program (NACO) trainer for the region and as part of the Minnesota NACO funnel. He retired from the Minnesota Historical Society after thirty-two years and then joined the staff at Minitex, University of Minnesota, in 2001, where he managed the contract cataloging service for nine years.

Edward was drawn to librarianship as a teenager, and his contributions to the larger profession started just as early. He joined the Minnesota Library Association (MLA) when he was still in high school and became active in the MLA Technical Services Section almost immediately, ultimately serving as president of MLA. He received the MLA President's Award in 1981 and also received an MLA Centennial Medal. Edward played a leadership role in the statewide shared integrated library system (MnSCU/PALS and MnLINK) Cataloging User Groups and Database Quality Maintenance Task Forces, where his expertise in authority control and indexing were particularly valued.

On a national level, Edward became a member of the American Library Association (ALA) in 1962. He served the Association for Library Collections and Technical Services (ALCTS) in a variety of roles. He was a member of the *Library Research and Technical Services* (*LRTS*) editorial board for fifteen years under various editors, ALCTS board member and parliamentarian, ALCTS International Relations Committee member, ALCTS Publications Committee member, and a member of many other committees and working groups. He indexed *LRTS* for decades, compiled the index for v. 1–25 in 1981; indexed the annual issues each year; and he compiled the cumulative index to v. 1–50. He served as the *LRTS* book review editor from 2003 until his death. He was a member and chair of the ALA Committee on Cataloging: Description and Access and MARBI Committee. Edward received the 2007 ALCTS Presidential Citation recognizing his lifetime of service to ALCTS.

Edward's contributions were not limited to the state and national level—he also was active in the International Federation of Library Associations and Institutions

as a member of the governing board and served on a number of cataloging-specific committees, including the Serials and other Continuing Resources Standing Committee.

Edward was the consummate cataloger and the epitome of a lifelong learner. He had a wonderful, dry sense of humor. His range of knowledge and willingness to share his expertise were extraordinary. We have lost a valued colleague and a dear friend.

Carla Urban, Edward's colleague at Minitex, observed, "Edward Swanson's career was characterized by a true love and understanding of cataloging; dedication to sharing that knowledge with others through training, one on-one consultations, and publication; and a commitment to the professional community and its activities. His generosity and dedication to colleagues and cataloging have been greatly appreciated."

Carla Dewey Urban (dewey002@umn.edu) is Coordinator, Minitex, University of Minnesota Libraries, Minneapolis.

Serials Literature Review 2008–9

Embracing a Culture of Openness

Maria Collins

The serials literature from 2008 and 2009 reveals the new identity of the serials professional—one who embraces openness. Many forces have pushed the serials profession into a state of flux; among these are the recent economic recession, the evolution of scholarly publishing, and the concept of open systems and data. Chaotic change for serialists has evolved into opportunities to revise collection strategies, approach Big Deal purchasing in new ways, devise creative user-access solutions, and become stakeholders in the debates over scholarly communication. The literature also reveals serials professionals developing a Web 2.0 sensibility. These themes are presented in the review through a discussion of six major topics: sustainability of serials pricing, the future of the Big Deal, management of electronic resources, access, blurring and decline of formats, and Web 2.0.

The financial crisis that began in late 2007, sometimes called the "Great Recession," sets the stage for the serials literature of 2008 and 2009. Wikipedia explains that this recession "is noted by many economists to be the worst financial crisis since the Great Depression of the 1930s."[1] With endowment funding at private universities dwindling and state universities receiving cuts and reversions, few library budgets have remained unaffected by the economic downturn. Using Wikipedia as the source for this quick explanatory blurb of the financial crisis is no coincidence—Web 2.0 functionality was another central theme pushed out to serials readers in 2008 and 2009. As serials professionals face reduced budgets, cancellations, and evolving publication models, Web 2.0 concepts of openness, interoperable systems, interactive communities, and networking are reshaping the manner in which people communicate, access scholarly content, and develop tools to support library collections. The Internet became the great equalizer—freeing content from traditional containers such as journal issues or books by providing a platform for distributing discrete units like articles and chapters. With growing support for the open access (OA) movement, evidenced by government- and university-supported mandates issued in 2008 and 2009, the scholarly communication process was officially transformed. Consequently, all participants of the scholarly journal information chain, including publishers, researchers, and librarians, are reinventing their roles in supporting scholarly communication. The serials literature reveals many ways in which the work of serials professionals and concepts of openness intersect: through experiments with acquisition models, development of standards to support open systems, automation and simplification of metadata to support patron access to electronic material, negotiation of rights to remove barriers to access, and an evolved focus on managing content instead of formats. Viewed holistically, these separate advancements coalesce to reveal a new global model for serials librarianship that prioritizes access, connectivity, and community.

Maria Collins (maria_collins@ncsu.edu) is Associate Head of Acquisitions, North Carolina State University Libraries, Raleigh, North Carolina.

Submitted July 14, 2010; tentatively accepted August 18, 2020, pending modest revision; revision submitted September 28, 2010, and accepted for publication.

I wish to acknowledge the contributions of my husband, Leonard Collins, for writing the reference notes and providing thoughtful edits to this paper (as well as looking after our three children). He has been wonderfully supportive and patient during the many hours spent on this project. I also would like to thank Liz Burnette, Head of Acquisitions at NCSU Libraries, for her continued support of my research and professional service.

This literature review provides readers with a snapshot of these themes through a sampling of periodical literature, occasional reports, and a small selection of books published in 2008 and 2009. The review covers six major sections: sustainability of serials pricing, the future of the Big Deal, managing electronic resources, access, blurring and decline of formats, and Web 2.0. topics included represent both the usual suspects within serials literature, such as serials pricing and access, and nontraditional topics pushed out to serialists from the serials literature, such as e-books and Web 2.0. OA and institutional repositories (IR) were also important topics during this period and appeared frequently in the serials literature. Because they are such large and important areas, they merit separate attention and will be the focus of a future literature review. To maintain a manageable scope for this paper, the author excluded many relevant topics, such as preservation, collection evaluation, open source, discovery, patron use of materials, the integrated library system (ILS) marketplace, and broader discussions of scholarly communication.

The author identified sources for the review through multiple means. The author initially selected for review journals used for past serials literature reviews as well as core serial titles that directly target serials professionals. Once these titles were identified, the author conducted a systematic table-of-contents analysis to identify quality articles and dominant themes in the literature for 2008 and 2009. Additional sources, including a select number of monographs and non–peer reviewed articles focusing on the article's key themes, were gathered organically through citations within the articles reviewed and additional literature searches of core themes in Library and Information Science Abstracts (LISA). In total, the author considered more than 350 articles, reports, white papers, and books. Each source selected was reviewed, abstracted, and assigned keywords and a quality ranking to assist with final selection and grouping for inclusion in the literature review.

Sustainability of Serials Pricing

The Economic Crisis and Sustainable Collections

Given the context of the economic recession, issues related to the serials pricing crisis were magnified in 2008 and 2009. The squeeze caused by increased serials prices and shrinking library budgets has forced many practitioners to reconceptualize sustainability of serials acquisitions through unbundling Big Deals and envisioning new models of access. Orsdel and Born capture this landscape well in their annual "Periodicals Price Survey" for 2008 and 2009 where they outline strong influences on the journal market, such as OA, the economic crisis, renegotiations of the Big Deal, and commercial pricing of journal content.[2]

To better understand the dominant pricing concerns for 2008 and 2009, summarizing a few of these key influences from Van Orsdel and Born's surveys is helpful. With the issuance of national and university mandates requiring research to be made freely available to the public (as directed by the National Institutes of Health (NIH), European Research Council, and many European and American universities, including Harvard), the OA movement transitioned from theory to practice.[3] Unfortunately, even with the increased push to comply with these policies, OA still had minimal affect on serials price increases. Van Orsdel and Born reported consistent increases of approximately 7–9 percent with the expectation of the same for 2010.[4] Even with strong pockets of opposition to these mandates as demonstrated by the Partnership for Research Integrity in Science and Medicine (PRISM), a lobbying organization pushing for the repeal of the NIH mandate, Van Orsdel and Born noted that publishers were not adverse to OA as a concept, but as a business model. Many publishers were simply scrambling to adjust their strategies for maintaining revenue streams while experimenting with OA models.[5]

Likewise, librarians were in the process of reevaluating their ability to meet increasing demands for online content from patrons while simultaneously dealing with decreasing budgets. This balancing act was precarious at best and was of such concern that several statements on the economic crisis were issued in 2009 in an attempt to emphasize to publishers the strained economic environments of consortia and academic libraries. Two statements to note include the International Coalition of Library Consortia's (ICOLC) "Statement on the Global Economic Crisis and Its Impact on Consortial Licenses" and the Association of Research Libraries' (ARL) "Statement to Scholarly Publishers on the Global Economic Crisis."[6] The ICOLC statement argued for the value of consortia in assisting libraries that are navigating economic strains on their budgets, saying that "library consortia are uniquely positioned to be the most effective and efficient means to preserve the customer base for publishers and create solutions that provide the greatest good for the greatest number."[7] The ARL statement provided additional feedback to publishers specific to academic research libraries. Both statements recommended that publishers allow flexible pricing and adjustments to negotiated deals for consortia licenses to stay in place whenever possible. The ARL statement further recommended that libraries be allowed to negotiate their contracts mid-term if necessary. Van Orsdel and Born noted "libraries and consortia have already begun invoking financial hardship clauses and asking to renegotiate licenses for bundled content midterm."[8]

Additional concerns in the ARL statement include the effect of the recession on a research library's ability to purchase "long tail" materials, such as foreign or small-press publications necessary to create a robust collection but

that often have a small circulation.[9] Often, these unique or niche collections are easily neglected in times of economic strain, yet these collections are becoming increasingly recognized as critical to a research library's value and mission. In a library environment where the Google Books program is sometimes perceived as homogenizing libraries' monographic collections, and commercial Big Deals that offer "everything we've got portfolios" appear to normalize journal collections across academic libraries, the argument can be made that long tail materials are the rarest and most valued of collections.[10] The question remains, though, whether libraries can build sustainable collections that include these types of materials in difficult economic times.

The topic of sustainable collections is a primary focus in Walter's article "Journal Prices, Book Acquisitions, and Sustainable College Library Collections."[11] Walters defined a sustainable collection as "one that can be maintained without significant degradation over time—one with a budget that provides for continued access to serial resources . . . as well as the timely acquisition of important monographic materials."[12] He discussed a radical suggestion for achieving a sustainable collection—abandon the serial (in bulk) for the monograph to support certain collections, such as an undergraduate collection. Walters argued that undergraduates often simply need exposure to topics that monographic resources can easily provide. Furthermore, a minimal number of serial resources are required to meet undergraduate research needs. To illustrate this point, Walters cited several studies that demonstrate "the most important research tends to be concentrated in a relatively small number of key journals."[13]

The serials literature contains many more examples of professionals using the crisis in scholarly communication as impetus to change directions and reconceptualize how to meet research needs in our current environment. The ARL statement invited publishers to take this same journey by modifying their business models and applying OA strategies as a possible transition from the traditional subscription model. Too often, business models for consortia deals penalize large ARL libraries that are forced to "absorb significant price increases to compensate for discounting other customers."[14] The concept of creating opportunity and change from hardship was also a prominent theme in Grogg and Ashmore's article, "The Art of the Deal: Negotiating in Times of Economic Stress."[15] These authors stated "times of economic stress can actually provide increased opportunities for negotiation."[16] For instance, if funding is short, libraries may decide to abandon ineffective tools that fail to support local workflows, or if a library is unable to sustain a consortia deal, perhaps a middle ground can be negotiated with a vendor that allows the library to reduce content to realize savings rather than abandon the deal altogether. Grogg and Ashmore wrote "the wisest among us will look upon the constricting economy as an opportunity to re-evaluate, renegotiate and revision."[17]

Understanding Journal Price Increases

To negotiate effectively, librarians must have a clear understanding of the factors that influence journal pricing. As information professionals consider the merits of an OA model, this understanding, especially of the commercial publishing market, becomes even more imperative. Ortelbach, Schulz, and Hagenhoff used a regression analysis to study several factors known from previous studies to influence journal pricing, including size of the journal, circulation, profit status of the publisher, country of origin, and academic discipline.[18] They found that profit status of the publisher and journal size were both positive factors influencing price. For-profit publishers produce more expensive journals than not-for-profit publishers, and journals with a more content are more expensive than journals with less content. However, the data were not significant enough to implicate these factors as causes of the serials crisis.

Another study by Kraus and Hansen found that journals from commercial publishers in chemistry were more expensive than those from noncommercial publishers.[19] They noted that "there is still a great discrepancy in cost between the commercial and non-commercial presses on a cost per journal, cost per downloaded article and cost per impact factor basis."[20] These are interesting findings in the context of the OA movement, which has the potential to revolutionize scholarly communication costs. Kraus and Hansen advocated submitting "articles to non-commercial and open access sources first, and publish in commercial journals as a last resort."[21]

In a presentation report for the 2007 North American Serials Interest Group (NASIG) Conference, Schonfeld described another concept that could be used to understand journal pricing: the two-sided market.[22] This theory has been applied to marketplace platforms with two or more sets of customers. The credit card market serving both merchants and the general public is an example. Schonfeld stated that "a scholarly journal also balances the interests of two sets of 'customers.' It must attract both sufficient articles to motivate scholars to read it as well as sufficient scholarly readers to motivate authors to submit articles to it."[23] This kind of analysis is especially useful with the advent of OA models causing a potential shift from the subscription model to the author-pays model.

Value and Role of Consortia in Journal Pricing

The value of buying through consortia during times of economic recession was another recurring topic in the literature

of 2008 and 2009. According to the *Survey of Academic and Research Library Journal Purchasing Practices* published by the Primary Research Group, 46 percent of the sampled libraries acquired subscriptions in bundles of fifty titles or more.[24] These bundled titles or packages were often procured through consortia negotiations. Perry surveyed ICOLC members to investigate current and future priorities for consortia.[25] Survey responses indicated that consortia continue to play a significant role in library acquisitions, especially in the context of the current economic downturn. Perry pointed out that the two current priorities, licensing and renegotiation, and the two future priorities, budget management and license negotiation, all relate to the economic downturn.

In their interviews with library professionals, Grogg and Ashmore also discussed a greater need to promote consortia value, especially state-level consortia. The authors pointed out the strong dependency many state-supported libraries have on consortia-provided collections that supplement their local collections.[26] Reduction of state funding for these consortia could have devastating effects on library budgets; librarians may find themselves in the position of renegotiating independently and allocating scarce funds for these unanticipated acquisitions. Several articles pointed forebodingly at South Carolina's dire situation.[27] Van Orsdel and Born noted "state funding for library consortia . . . tumbled in a number of states—South Carolina's PASCAL lost 90 percent of its funding."[28]

The ARL statement also advocated for the role of consortia negotiations but tempered its response with caution about unsustainable models.[29] Large research libraries are becoming unable to subsidize consortia agreements by taking on an inordinate amount of consortia costs for the Big Deal. Alternative pricing models that provide a more equitable distribution of financial costs for consortia arrangements are needed for the future viability of these deals.

Another article in Grogg and Ashmore's "The Art of the Deal" series focused on advantages and disadvantages of consortia negotiations.[30] The authors noted the obvious benefits of bulk purchasing and obtaining more content for the money; they also listed the lowered costs and redistributed savings, negotiation support of individual librarians, and consortia influence to include more "progressive licensing terms."[31] Regarding disadvantages, they discussed the often slow pace of consortia negotiations, losing the ability to make flexible collection decisions, complexity of negotiations with a large number of members, and potential consequences of loss of access when consortia members opt out of consortia agreements.

Sanville added to this discussion of the role of consortia with a frank assessment of the economic value of these deals.[32] Consortia deals have proven their worth by increasing access and reducing costs, but Sanville also noted that consortial buying has been a treatment of the symptoms created by the crisis in scholarly communication, but is not the cure. In this practical context, Sanville listed several issues not addressed through consortia deals, such as increased production of scholarly content, the continued prevalence of traditional publishing models, smaller publishers being forced to increase prices to keep up or merge with larger publishers, and increased patron demand for online content. As scholarly communication evolves, the process of consortial buying also may change as these concerns are addressed.

The Future of the Big Deal

Parallel to these discussions of consortia value was concern about the sustainability of the Big Deal or all-inclusive publisher packages often negotiated by consortia for a minimal fee on top of a library's historic expenditures with that publisher. Numerous articles touched on a variety of issues related to the Big Deal, including a special section in the 2009 volume of *Serials Librarian* edited by David Fowler. Discussions of advantages and disadvantages of the Big Deal were the most common threads.[33] A summary of the most commonly mentioned advantages includes special pricing with lower unit prices; increased access, especially for smaller institutions; controlled price increases; immediate access to content; and streamlined workflows. With Big Deals often consuming a large portion of libraries' materials budgets, libraries also experienced disadvantages noted in the literature, including an inability to manage local collections and reduced flexibility to support alternative access models.

The volume of articles on this topic is indicative of the tension concerning the sustainability of the Big Deal. Grogg and Ashmore even speculated "time will tell if the world economic crisis of 2008, 2009 and beyond will finally be the straw that breaks the big deal's back."[34] Torbert asserted that satisfaction with the Big Deal is decreasing, but librarians surveyed from academic libraries still believe that Big Deal "benefits outweigh the difficulties."[35]

Taylor-Roe's discussion of her frustrations with the Big Deal also illustrates concerns about sustainability.[36] She detailed results from a survey of United Kingdom libraries concerning their satisfaction with the Big Deal. Specific discussion points included a dislike of Big Deal business models that limit cancellations, base pricing off a library's historic expenditures, and inhibit a library's ability to purchase journal content from other publishers. Rolnik, a small publisher, expounded on this last point by stating the Big Deal "locks my content out of the budget. There is often little budget remaining after the library pays all of its Big Deal invoices, even for high value content."[37]

Other frustrations mentioned by Taylor-Roe, such as difficulties in tracking transfer titles and an inability to deliver to

libraries an accurate title list, related to a publisher's ability to provide quality customer service.[38] Given these frustrations in combination with the economic downturn, the 14 percent of respondents to Taylor-Roe's survey planning to cut back on their Big Deals is not surprising; 38 percent thought they would maintain the Big Deals they already had, but a striking 28 percent indicated they would only take out new deals if they could cancel old ones, and 24 percent said they would actively seek to reduce Big Deal expenditures. Based on these comments, a breaking point is coming. Flexible pricing and choice of content to reduce costs are essential elements needed for a sustainable Big Deal model.

Usage Studies of the Big Deal

One method to evaluate whether a library is ready to break their journal packages is to study use of the collection. One criticism directed at the Big Deal is the prevalence of nonused titles within the package. Several studies examined use of their titles to determine whether this concern pertained to them. Botero, Carrico, and Tennant conducted an evaluation of Big Deal packages for the University of Florida libraries using Counting Online Usage of Networked Electronic Resources (COUNTER) compliant, publisher provided usage statistics.[39] For this library, the low cost per article and increased access to content indicated a deal too good to refuse despite concerns about the collections budget. Another study by Termens noted that journal package use across institutions within a consortium is not equivalent.[40] Termens' research indicates that the level of use may not always be high or even predictable. One final study of note is Murphy's analysis of nonused titles accessed through the OhioLink consortium.[41] Murphy conducted a citation analysis to understand the use of titles within select departments and discovered that fewer than 50 percent of the titles cited by faculty were purchased as part of Big Deal packages, and 75.4 percent of the titles maintained through OhioLink were cited minimally by faculty during a three-year period. Murphy argued that the premium paid to provide access to these unused titles should be funneled elsewhere to fund quality content requested by faculty. She aptly stated that "rather than contributing to the normalizing of library collections by supporting the strategic positioning of commercial publishers, large research libraries may respond better to the needs of all faculty, especially those conducting research in smaller fields, by returning to a la carte purchasing."[42]

Reinventing the Big Deal

Numerous articles explored alternatives to the Big Deal that might better serve as a sustainable option. Van Orsdel and Born commented that in reaction to the increasingly resource-intensive nature of the negotiations required to set up Big Deals, "some commercial publishers are talking about getting out of the negotiating business and are considering selling their journals as a single database with fixed prices."[43] Best described the concept of an orderly retreat from the Big Deal whereby libraries negotiate "the right to cancel superfluous or unaffordable titles."[44] Cleary supported this option noting that this would allow members to stay in the deal, provide access to the highest used titles, and reduce costs to a manageable level.[45] Another reconceptualization of the Big Deal incorporated aspects of both of these concepts. Boissy called his new version of the Big Deal the "comprehensive consortial model."[46] In this model, publishers grant archival rights, "streamline the consortial model, find a sustainable pricing point, and grant full rights to the complete publisher journal portfolio in electronic form."[47] One more option is to unbundle the Big Deal, not wholesale, but deal by deal depending on an understanding of return on investment. Cleary described a useful example of the unbundling process undertaken by the Queensland University of Technology (QUL) for their Taylor and Francis packages.[48] For the three Taylor and Francis collections purchased by QUL, Cleary conducted a cost per full-text-download analysis, revealing QUL's return on investment for these collections. Librarians for QUL established a benchmark of acceptable full-text downloads as an indicator of sufficient use to assist in their assessment. Ultimately, the decision was made to unbundle one of the three collections and return to a la carte purchasing for those titles. Cleary argued that consortia should negotiate to remove low use and low value titles to reduce costs or face cancellation of the Big Deal.

Mitchell and Lorbeer provided a well-written case study of a library transforming its collection after cancelling a Big Deal.[49] For them, the issue of nonused titles made continuing their Big Deals untenable. Purchasing larger numbers of nonused titles simply did not equate to their idea of an economically sustainable collection. Luckily for the University of Alabama at Birmingham (UAB), the library has a successful liaison program and, armed with COUNTER stats, librarians were able to transform the collection by selecting high-use titles, greater use of interlibrary loan, and increased support for pay-per-view. Most notably, they have been able to accomplish these changes with minimal disruption to patron service.

Pay-Per-View as an Alternative Model to Subscriptions and Big Deals

Outside of discussions about OA in the serials literature, pay-per-view (PPV) appeared to be the alternative model of choice to the Big Deal. Two works—by Carr and by Harwood and Prior—provided thorough examinations of PPV, including discussions of models, advantages and

limitations of the service, and the future of PPV.[50] As defined by Carr, PPV in simplest terms is a purchasing alternative for journal content in which the "library acquires individual articles that users request."[51] PPV models can take numerous shapes and forms. Harwood and Prior's descriptions of usage-based e-journal purchasing discussed two models trialed by the Journals Working Group (JWG) of the United Kingdom's Joint Information Systems Committee (JISC): (1) a "pay-per-view converting to subscription model" whereby patrons purchase articles up to a set threshold, at which point the journal title converts to a library subscription with unlimited access to articles, and (2) the "core plus peripheral" model whereby a publisher offers a subscription to core titles and PPV access to nonsubscribed titles.[52] Carr also outlined varying implementations of PPV from six different academic libraries in which the library established an account with a content provider, users initiated the purchase, and the library subsidized the costs.

The driving force behind most of the implementations was the realization of cost savings while maintaining a similar level of access as previously experienced with the Big Deal.[53] In a workshop report by Wolverton, additional advantages of PPV were outlined by Tim Bucknall, who implemented the service at the University of North Carolina at Greensboro in 2002.[54] These advantages include increased availability of journals and backfiles, immediate access to articles, a more cost-effective model to support access for low-use journals, and benefits as a collection development tool.

These works also discussed the limitations of PPV. The trials implemented by the JWG revealed numerous findings that could prove challenging as PPV models evolve.[55] One technical concern involved excluding freely available content to calculate chargeable downloads. To make these calculations, content providers need to devise ways to filter out freely available content, such as OA articles and promotional content. Usage statistics from multiple models of access (PPV, subscriptions, and packages) also would need to be consolidated for a complete understanding of use. Budgeting for PPV is another source of concern; dependable projections are needed to stabilize or account for PPV expenditures. These findings reveal that administrative support could be substantial, especially for monitoring usage statistics. Harwood and Prior also noted that archival rights often are not provided for PPV-provided titles unless they are converted to a subscription model. Collection managers would need to consider PPV options carefully if ownership of content is a strong priority for their organization.

Carr responded to and expanded on many of these concerns.[56] Regarding administrative support, respondents to his survey noted the level of support to be similar to managing subscriptions and packages. In addition, survey respondents noted that they did not experience a "high degree of uncertainty or risk" in respect to managing their budgets; instead they realized cost savings by using this model.[57] In a short essay for *Against the Grain*, Carr also expanded on the issue of archival and perpetual access rights to PPV purchased content.[58] Essentially, a library's decision and commitment to a PPV service comes down to their readiness to reconceptualize the role of a library as a keeper of the keys to owned library content. In the current environment, with an increasing critical mass of materials available that the library cannot afford and an increasing emphasis on the openness of content, the library's role as purchasing agent for scholarly materials is slipping away to be replaced by a browser and the Internet. . . . Carr suggested that perhaps now is the time for libraries to commit to a "just in time" philosophy to meet the needs of users. He stated "many libraries today are in fail-fail situations. Librarians might reason that it is better to face the possibility of failing anticipated patrons in the future than the certainty of failing real patrons in the present."[59]

What is the future of PPV? Will it replace the Big Deal? Most of the libraries surveyed by Carr supported PPV.[60] Carr did note publisher acceptance of PPV as an ongoing issue. Two case studies—one presented by Chamberlain and MacAlpine, the other by Wolverton and Bucknall—also showed positive PPV implementation results.[61] These studies implied that PPV has a future in a mixed-model purchasing environment. The investigations of the JWG resulted in a more cautious view of PPV. Harwood and Prior made the following assessment on the future of PPV: "It was felt that on the basis of findings from these trials, a 'traditional' Big Deal pricing model is likely to give much greater budgeting predictability, whilst still offering access to all titles from the participating publisher."[62]

Managing Electronic Resources

For many serials professionals, 2008 and 2009 were the beginnings of a second decade of managing serial resources online. Evolving management practices and technological solutions for effective electronic resources management (ERM) continued to be primary topics in the literature. However, more abstract and global themes, such as strategic planning, interactive systems that support the automatic exchange of data, and the erosion of the journal issue as the dominant format or unit supporting the distribution of research also were discussed. Given the context of an open scholarly communication system and the influences of the Internet, this evolution in the research describing ERM is not surprising. Many information professionals have experienced the challenges and frustrations of developing or implementing an ERM system. Through this process they have gained an evolved understanding of the need for standards to support and enhance ERM system functionality. The implementation of ERM tools by libraries and the often

painful experience of modifying and abandoning traditional workflows have taught many librarians the value of careful planning and establishing transparent communication strategies. Finally, with more sophisticated systems in place, serials professionals also are seeing the boundaries between public and technical services disappear as technical services staff provide frontline support to patrons to ensure seamless access for electronic resources. To capture these themes, this section will discuss the following ERM-related topics: planning and workflows, ERM tools, and evolving standards and best practices.

Planning and Workflow Analysis

Abstract concepts like planning and workflow need to be grounded in both theory and practice to gain the true understanding of the reader. Several articles that focused on planning and workflow analysis for ERM provided this conceptual range from abstract to practical. Two articles took a predominantly theoretical approach in their discussion of electronic resource planning. The first, by Hulseberg and Monson, provided an interesting case study outlining Gustavus Adolphus College's strategic planning for ERM.[63] After initial brainstorming and goal-setting, librarians at the college established several initiatives related to ERM, including analyzing workflows, creating a documentation system, and developing an ERM system (ERMS). To evaluate the successful achievement of these initiatives, they determined appropriate assessment tools for assistance. Examples include usage statistics, library faculty and staff surveys, an annual workflow analysis project, and an annual report on recent research. Planning for ERM in this strategic fashion was valuable in enhancing their management process.

An article by Collins discussed theoretical concepts of planning and workflows with an emphasis on the importance of effective communication.[64] The author noted that "e-resource management concerns have outgrown traditional department boundaries necessitating efficient communication strategies to stabilize and guide workflow practices across the library."[65] The article provided comments from eight librarians about workflows given up and maintained after the advent of electronic journals in their libraries. Many of the processes left behind were print-based, including claiming, check-in, and binding. Librarians also were decreasing serials cataloging and increasing reliance on MARC record services. Processes retained included maintaining relationships with agents, verifying online access, and creating access points for all e-resources in the catalog.

In terms of practical approaches for workflow analysis, Blake and Stalberg provided a well-written explanation of a method to observe e-resource workflows in which Blake shadowed each staff member in her department for an entire day.[66] The authors transcribed observations into a visual representation of the serials life cycle and gave technical services staff an opportunity to respond to these charts and provide feedback. In addition to gaining a better understanding of serials workflows both in and outside the department, this analysis resulted in enhanced communication between serials staff. This process assisted staff in understanding changes in practice because of an increased focus on ERM, helped staff in defining responsibilities, identified legacy practices that could be abandoned, and revealed miscommunications and inconsistencies in the workflow that needed attention.

Strader, Roth, and Boissy provided a practical application of workflow analysis in their discussion of roles and responsibilities of parties involved in the information supply chain.[67] They presented a checklist of tasks needed to provide access to an electronic journal, and they identified the responsible party (publisher, agent, or librarian) for each step. For each task, the chart notes the appropriate communication needed from each party and describes the processes each party must complete to establish and maintain e-journal access. This is an excellent example of a workflow document that can be used to organize and direct those who support electronic journal access.

Electronic Research Management Tools

Support for ERM workflows often comes in the form of an ERMS, including commercial and locally developed tools used to facilitate ERM. These options are abundant and of strong interest to serials professionals, as demonstrated by the prominence of this topic in the literature. Marketplace overviews are common and assist librarians in analyzing the best choice for meeting their ERM needs. The "Helping You Buy" series in *Computers in Libraries* always provides useful information for considering a technological purchase. Breeding provided the latest edition of this column focusing on ERMS, reviewing six commercial products.[68] Breeding defined these systems and presented areas of the ERM life cycle that an ERMS can support. He also provided a useful comparison of each system's knowledge base, or central data repository, detailing a library's collections. Other areas mentioned included local installation versus software as a service (SaaS), integration with the ILS, license management, statistics, and reports. Another article by Collins provided an exhaustive review of nine ERMSs, including one open source system.[69] Collins provided feedback from the librarian community about the challenges of ERM, including change management, prioritization of work, and inconsistent practices and metadata. Representatives from the companies interviewed also discussed challenges they have experienced in building an ERMS, including the lack of industry standards and interoperation with other systems.

The librarians' wish list for ERM functionality included automated and custom reporting, the ability to manipulate data within the system, continued need to minimize duplicate entry, and additional management features to support e-books, complex workflows, and troubleshooting.

Other discussions of ERMS in the literature focus on locally developed or open source options. Stranack discussed the CUFTS ERM in his presentation of the reSearcher software suite, which is an open source software option from Simon Fraser University to support link resolution, federated searching, personal citation management, and ERM.[70] Stranack promoted the benefits of collaboration within an open source community and invited readers to consider this alternative to the "high cost and closed nature of commercial software."[71] ERMes was another open source solution depicted by Doering and Chilton.[72] Disappointed by commercial offerings, librarians at the University of Wisconsin began locally developing their ERM using an access database. This allowed them to create a quick, simple, yet functional solution to meet their ERM needs.[73]

Many libraries are using simpler tools and solutions to facilitate ERM processes and are finding that these alternatives are quite effective in meeting their needs. A presentation report from the 2007 NASIG conference outlined alternatives including FileMaker Pro, Blackboard, and EBSCOhost's Electronic Journal Service.[74] Watson and Hawthorne used these alternatives in lieu of commercial systems at their respective institutions to support the storage of ERM metadata, management of invoices and license agreements, registration tracking, and follow up including ticklers, notifications, and alerts. Murray, dean of libraries at Murray State University, described Web 2.0 solutions as an alternative to commercial software.[75] Librarians at Murray State used blogs to assist with trials, Google Docs and spreadsheets to support subscription management (cancellations, title changes, and new orders), wikis for storage of administrative metadata, widgets for reporting, and mashups to bring these technologies together in a single interface. Murray commented that "commercial ERM systems may seem at times to have too much functionality to be practical, particularly for libraries that do not need all the bells and whistles."[76]

Many ERMS users have quickly discovered that their needs for technical support go far beyond the idea of a centralized container to store metadata and a system to support license management. Libraries need ERM solutions that actively support and manage ERM workflows and capture and interlink complex relationships between organizations, resources, collections, and local management data needed to categorize their electronic resources. A few locally developed alternatives discussed in the literature have created custom solutions to address some of these iterative or complex issues. Collins and Murray described the University of Georgia's (UGA) electronic journal verification system, which automates the scheduling, queuing, and problem-reporting pieces of an access verification workflow.[77] While recognizing the need to verify electronic subscriptions, many librarians have not been able to adopt a proactive approach to access verification because of the time-intensive nature of access checking. UGA's system makes this process more viable and serves as a potential solution for this particular ERM challenge.

Resnick and colleagues also described a creative method to assist with access problems.[78] Their paper recounted the evolution and development of a problem-reporting help desk database. The creation of this tool was the culmination of an experiment to include technical services librarians with ERM and licensing expertise as part of the help desk service. Including these librarians resulted in improved response time and problem resolution for access problems and eventually led the team to develop a helpdesk database to support the problem-resolution process.

Blake and Samples described a local solution used to resolve issues surrounding metadata relationships in ERM systems, specifically, organization name authority.[79] Their article described the implementation of a name authority tool as part of North Carolina State University Libraries' locally developed ERMS, E-Matrix. The authors provided context for the name authority problem through interviews with librarians across the country and discussed current practices of organizations such as OCLC, which also support name authority data.

The success of these locally derived solutions comes from a strategic focus on a given problem. These efforts contribute to a greater understanding of pieces of the ERM dilemma and, if considered more broadly, could lead to more effective ERM designs for commercial products. These local developments also reflect the good will of the librarian community as many of the institutions designing these solutions are willing to share the metadata or code behind these tools.

Going beyond descriptions of ERMSs, the serials literature also discussed their implementation. Case studies by White and Sanders as well as Beals provided a detailed description of the investigation, selection, and implementation of an ERMS.[80] Other articles provided tips for implementation success.[81] These include the importance of teamwork and communication; allocating the appropriate amount of time and staffing resources; documenting e-resource workflows; setting goals and priorities before beginning implementation; determining local programming resources; matching collection size and complexity to ERMS functionality; establishing target date and deadlines for implementation phases; and employing change management strategies, such as regular meetings, additional training, and inclusion of affected staff in the planning process.

Several of these tips were lessons learned from an unsuccessful or challenging implementation. Grogg described

the often slow pace of implementation in her column "Electronic Resource Management Systems in Practice," and this is demonstrated by Ekart's description of Kansas State University's struggles to implement Verde.[82] Numerous issues hampered the speed of implementation including defining workflows in Verde that would account for local practices. Ekart described their workflow implementation attempts as "trying to shoehorn our current processes into the available site-specific tasks."[83] Grogg echoed these concerns, stating that other librarians have reported frustrations with an ERMS's ability to handle complex workflows.

Another troubled implementation story from Pan described one without an initial defined purpose or goal.[84] Because of miscommunications with their vendor, the staff at Pan's library began with efforts to display their ERMS's resource records in the catalog. After spending a year designing a workflow and troubleshooting this process, library staff abandoned using their ERMS for public display and reevaluated their purpose for implementation. They masked the ERMS records in their system and adjusted their focus to support backend management data for acquisitions, licensing, and collection management. This library learned the difficult lesson of appropriately aligning processes with priorities to successfully meet the institution's implementation goals. Of course, implementation goals often vary across institutions. The most common ERM goals mentioned throughout these implementation case studies include centralizing ERM data, supporting the subscription life cycle (trials, renewals, new subscriptions, and cancellations), supporting collection analysis and storage of usage statistics, limiting multiple points of data entry, managing license information, and streamlining workflows to facilitate patron access.[85]

The serials literature reveals the complexities of ERM as well as the evolution of tools to address those complexities. The implementation experience has taught librarians the value of planning to help define the purpose of an ERMS and the importance of matching a system's functionality to this purpose. Given the extensive research in this area, change to "ERMS" appear to be a desirable component of ERM. One final article of note edited by Tijerina and King provided five essays exploring the future of ERMS.[86] The common thread through all these essays is the importance of standards in the continued development of ERMS. Riggio commented that vendors are currently moving in the right direction with the development of standards like the Standardized Usage Statistics Harvesting Initiative (SUSHI), Cost of Resource Exchange (CORE) and the ONline Information eXchange (ONIX) family. Systematic use of standards to normalize and support metadata transmission will ultimately lead to more open ERMS. In another essay from this article, Pesch reiterated this theme, commenting that ERMS cannot survive as standalone systems. He argued that "ERM systems must become a part of the e-resource supply chain."[87] This is only possible if these systems use standards that allow for the automation of data exchange, including knowledge base holdings, cost data, and license data. Pesch commented further that standards are critical in reducing the effort currently required to maintain and populate an ERMS.

Standards and Best Practices

Given the previous statements about the prominent role standards will play in the future development of ERMS, the topic of standards takes on a new dimension. The serials literature for 2008 and 2009 provided updates and explanations of new and evolving standards, best practices, and projects in use or development that support ERM. Reasons for many of the standards and initiatives mentioned can be grouped into three areas. Initiatives like the ISSN-L and Knowledge Bases and Related Tools (KBART) facilitated journal linking and interoperability across systems. COUNTER, SUSHI, and CORE supported the structure and protocol for transferring usage and cost data to support cost-per-use. Finally, initiatives like Shared Electronic Resource Understanding (SERU) and Project Transfer (www.uksg.org/transfer) addressed librarian frustrations with aspects of ERM, such as licensing and the transfer of titles across publishers. Numerous articles provided basic updates to standards, and the few selected for the review provide useful overviews.

As part of his standards column, Pesch provided a succinct explanation of the ISSN-L introduced in August 2008, which serves as both a title-level and medium-level identifier.[88] Pesch included easy-to-understand graphics that illustrate the failings of the current ISSN and the successes of the new ISSN-L to support linking across systems. In addition, Pesch discussed how the ISSN-L standard can be implemented across the industry by using ISSN mapping tables for the already assigned ISSN-Ls. He explained that all participants in the journal information chain will need to implement this standard for it to be successful, but once in place, the ISSN-L "should result in significant improvements in the quality of linked access."[89] Vincent, from the ISSN International Centre, further defined the need for the ISSN-L standard as twofold: ISSN users want a way to identify a "product (or manifestation) level," and a standard is needed that will "collocate . . . medium-specific versions" of a title.[90] Vincent explained that the first ISSN assigned to a title, no matter the format, will be designated as the ISSN-L. She also provided an explanation of how MARC fields will accommodate this new data element with the addition of subfield "l" in the 022 ISSN field.

Another initiative, KBART, which is a joint National Information Standards Organization (NISO) and United Kingdom Serials Group (UKSG) project, also proposed to enhance linking through cleaner metadata. A presentation by McCracken and recorded by Arthur explained that the

goal of KBART is to improve "the functioning of OpenURL by providing standards for the quality and timeliness of data provided by publishers to knowledgebases."[91] If the metadata supplied to knowledge bases by content providers were improved, many OpenURL data and syntax errors would be resolved, allowing for more seamless linking to electronic resources. Currently, the project is focusing on best practices for publishers, but McCracken stated that this initiative could lead to future standards for publisher submission of metadata to industry knowledge bases.

Other standards address some aspect of the cost-per-use equation, a highly desired metric that many librarians would like to generate from their ERMS to facilitate collection management. The first element of cost is difficult to collect for cost-per-use given the myriad pricing models, billing bundles, and package deals that librarians manage in acquiring electronic resources. In addition, much of these cost data are locked within an ILS and not easily extracted. Rather than perform duplicate data entry to capture cost data in an ERMS, practitioners are investigating means to easily extract cost information from an ILS and transfer it to an ERMS. The *White Paper on Interoperability between Acquisitions Modules of Integrated Library Systems and Electronic Resources Management Systems*, written by a subgroup of the Digital Library Federation's Electronic Resources and Management Initiative, explored this issue.[92] Through an evaluation of four case studies, the authors were able to identify seven acquisition-specific data elements necessary to transfer between systems to support cost-per-use. These elements included "purchase order number, price, start and end dates for the subscription period, vendor name, vendor ID, fund code, and invoice number."[93] In addition, the paper discussed the various challenges of cost-per-use data, including the lack of itemized pricing for many journal packages and the difficulty in valuing unsubscribed titles that are often part of consortia Big Deals. Because of this white paper, the Cost of Resource Exchange (CORE) standard was initiated. Riding provided an explanation of the development of the CORE standard in "Cost of Resource Exchange (CORE): The Making of a Library Standard."[94] A NISO working group was formed, and an XML schema was written to test CORE as a draft standard from April 2009 through March 2010. Riding commented that the realized value of the standard will come when "librarians are able to use it to pull cost information from the ILS and other systems into the ERMS to make intelligent renewal and purchasing decisions."[95] This will be a highly anticipated moment for librarians desiring cost-per-use functionality in their ERMS.

Use of electronic resources, the other component of cost-per-use, has received dedicated treatment over the last few years through continued development of the COUNTER Code for Practice and SUSHI standard. Much of the focus in the literature concerning these standards centered on continued challenges of collecting use data. In Pesch's update on the COUNTER Code for Practice in release 3, he said the most important addition was the requirement that content providers support SUSHI to be COUNTER-compliant.[96] Pesch hoped this would be a huge step forward in establishing large-scale support for SUSHI.

Other additions to the COUNTER Code of Practice relate to a provider's responsibilities to separate specific use activities that skew usage statistics. These activities include federated search sessions, pre-fetch activity, robot activity, and the retrieval of full text for archiving. All these activities need to be excluded or identified for COUNTER compliance. Matthews further expanded on these challenges in her overview of the advances in usage statistics.[97] She commented that many vendors, especially smaller operations with limited development resources, may have trouble complying with some of the release 3 additions without a major overhaul of their systems. To address these kinds of limitations, many vendors are utilizing third party intermediaries to serve as a platform for content delivery and to capture usage data.

Gedye explored measuring the use of individual research articles through a detailed explanation of the Publisher and Institutional Repository Usage Statistics (PIRUS) project.[98] He noted increasing interest in gathering usage statistics at the article level, especially because government and university policies mandate self-archiving in institutional repositories. He commented that this kind of data can serve as an additional measure of article-level and journal-level quality. Furthermore, a standard is needed to consolidate article-level use from multiple sources, such as publisher-controlled platforms and locally maintained IRs. Participants in the PIRUS project hope to develop COUNTER-complaint usage reports at the article level, create guidelines to assist any host of online journal content to create these reports, and suggest a method for report consolidation.

Other areas of standards development are focusing on best practices to simplify ERM processes. NISO's adoption of SERU in 2008 as a best practice represented an attempt to simplify the time-consuming and complicated license negotiations that are often necessary to acquire electronic resources.[99] According to NISO, "SERU offers publishers and libraries the opportunity to save both the time and the costs associated with a negotiated and signed license agreement by agreeing to operate within a framework of shared understanding and good faith."[100] The SERU document provides an introduction to the concepts behind SERU, guidelines for implementation, and a statement of common understandings for subscribing to electronic resources. An article by Chamberlain and a NASIG conference report by Smith both provided background and context for SERU as well as suggestions for implementation.[101] Both parties involved in the purchase agree not to license the resource but abide by SERU and copyright law; they then sign the

registry on the SERU site and include any special terms concerning the arrangement in the purchase order.[102] Chamberlain commented further that SERU is not intended to replace all license negotiations, but when both parties are in agreement and comfortable with a simple understanding, this process can greatly streamline negotiations.

Another initiative—Project Transfer, sponsored by the UKSG—defines voluntary best practices for transferring a journal title between publishers. Pentz and Cole provided the most recent update of Project Transfer during 2008–9 in their paper "The UKSG TRANSFER Project."[103] They explained that Project Transfer will ensure that transferred content remains accessible and that users will experience minimal disruption. In addition, established best practices should create a framework to support more efficient processes for transfers and create expectations of the roles of each of the parties involved. Pentz and Cole also provided a detailed overview of the 2.0 version of the code of practice. They indicated that this latest version has received more publisher support than earlier versions, as twenty-five publishers had endorsed the code by May 2009.

Books and ERM: Planning, Change Management, and Practice

The breadth of topics related to e-resources addressed in the serials periodical literature is indicative of the fundamental changes occurring in serials management and practice. Over the last decade, experimental practices in managing transitional collections as well as developing and integrating ERM tools within both the library and global information environment have resulted in the formation of ERM as a discipline of study and practice. Even though this review focuses primarily on periodical literature, a small but significant number of monographs focusing on a wide-range of ERM topics should be mentioned. What follows is a list and brief description of these monographic resources.

- Sheila S. Intner with Peggy Johnson, *Fundamentals of Technical Services Management*.[104] As technical services departments evolve to handle the changes in acquisition, cataloging, and preservation practices, due partly to the addition of electronic and digital resources in library collections, new managers to technical services need practice guidance and advice to strategically plan for and manage these changes. This book provides management theory, tips, and additional reading suggestions concerning the role of the technical services manager, staffing practices, evaluating a technical services department, understanding and managing the impact of digital resources in the department, and maintaining vendor relationships.
- Peter M. Webster, *Managing Electronic Resources: New and Changing Roles for Libraries*.[105] This book provides an extensive overview of ERM within the context of the larger, integrated information environment. Webster provides a useful and easy to understand presentation of technical ERM initiatives and tools such as link resolvers, citation managers, ERM tools, social networking applications, and ERM-related functions within the ILS such as link checking, package management, and authentication.
- Maria D. D. Collins and Patrick L. Carr, eds., *Managing the Transition from Print to Electronic Journals and Resources: A Guide for Library and Information Professionals*.[106] Electronic resources have transformed library collections, workflows, staffing practices, patron interactions, and management tools. This monograph includes a wide range of topics focused on these transitions, including budgeting and acquisitions, criteria for selection, collaborative library-wide partnerships, institutional repositories, ERMS, data standards, workflow management, and e-resource licensing.
- Rebecca S. Albitz, *Licensing and Managing Electronic Resources*.[107] Albitz provides a discussion of electronic resources licensing and library rights within the context of copyright law. This book also discusses the terms and conditions within a license, model agreements, licensing alternatives, and best practices for license negotiations.
- Holly Yu and Scott Breivold, eds., *Electronic Resource Management in Libraries: Research and Practice*.[108] Planning and workflow management are a central focus of this reference source with an emphasis on the electronic resource lifecycle. Management practices are provided for electronic resource selection, acquisitions, cataloging, public display, and usage evaluation.

Access

Simplifying the Rules

The acquisition of thousands of electronic journals and e-books by many academic libraries in a short period combined with increasing patron demand for electronic content has resulted in a renewed focus on quick, efficient access to materials. Initiatives described in the literature reflected this emphasis on access through discussions of metadata simplification and process automation. Several articles focused on the revisions of cataloging rules or standard records to simplify the level of metadata required. A study by Terrill discussed the new CONSER Standard Record, which "represents a change in cataloging philosophy, with its emphasis on

access and meeting user needs over detailed description with its focus on access rather than bibliographic description."[109] Terrill hoped to assess the initial acceptance of the CONSER Standard Record by catalogers and determined that library staff in the study had accepted most of the changes with minimal modifications to the records during copy cataloging. When individual MARC fields were considered, cataloging staff, especially those from CONSER libraries, were unlikely to edit individual fields in 68–99 percent of the instances. With a minimalist approach to description, Terrill indicated that the new CONSER Standard Record should lead to "more efficient and less expansive cataloging."[110]

Kemp examined numerous recent developments in cataloging, many of which will simplify cataloging practice, including the CONSER Standard Serial Record and the Resource Description and Access (RDA) cataloging rules.[111] Regarding the CONSER Standard Serial Record, Kemp stated that this record should consist "of common elements that could apply to any serials title, print or online, with just one level of detail, rather than allowing several different levels. The new standard record would provide all the basic information necessary to allow users to differentiate between, collocate, or find desired titles."[112] She commented further that the cataloging rule revisions behind RDA also should allow for more flexible cataloging of electronic resources. For instance, the Functional Requirements for Bibliographic Records (FRBR) throughout RDA will support separate descriptions of the content and the container delivering the content. Concerning RDA and the CONSER Standard Serial Record, Kemp stated that "not only will the rules of how to input information into MARC tags become simpler and easier to find, but the most detailed level of cataloging for serials will become less complicated."[113] Another source for information about the treatment of serials using RDA is Curran's column "Serials in RDA: A Starter's Tour and Kit," in which she identified the major sections of RDA that apply to serials.[114] She outlined changes in the cataloging rules from the *Anglo-American Cataloging Rules*, 2nd ed., as they relate to serials cataloging, and discussed specific rules related to serials from the November 2008 RDA draft. Serial issues requiring attention after the first release and RDA and FRBR mappings also are included. For serials catalogers interested in learning how RDA will affect their work, this RDA primer will serve as a useful cheat sheet to serials-related rule changes.

Value of Automation

In today's library environment, management of access is no longer just the domain of the bibliographic record; this process also occurs through title activation and management in a library's knowledge base. The catalog and the knowledge base also have become increasingly intertwined because of the use of a knowledge base to manage MARC record services to automate the process of cataloging. In an information world where hundreds of serial titles can be acquired through a single purchase, use of MARC records services and vendor-created MARC data has become critical to a library's ability to provide quick access to these collections. Kemp commented that outsourcing of cataloging through MARC records sets and services has allowed libraries to reduce the labor involved in serials cataloging and refocus staffing resources on original cataloging.[115]

In a similar vein, Chen and Wynn stated that "it is not possible to provide access in the library's catalog to all of these e-journals through manual cataloging alone."[116] Results from their survey of academic librarians across the United States concerning e-journal cataloging practices indicated that manual cataloging still occurs, but less frequently, as libraries work to automate serials cataloging through purchased MARC record sets. Chen and Wynn also stated that several libraries regarded "e-journal cataloging to be an 'unnecessary luxury' or even a waste of both time and resources."[117] This kind of statement represents a philosophical shift not only concerning how serials cataloging should be performed but also for the role of the catalog in providing access to journals. Chen and Wynn continued by reporting that "a growing number of libraries no longer consider the library catalog to be the primary means of access to e-journals. Instead, they direct users to tools other than the catalog for finding them."[118] An article by Lowe provided an excellent example of the shift away from the catalog as the primary discovery tool for electronic journals.[119] The library in Lowe's article added all of their print holdings to their knowledge base, which already tracked their electronic holdings, to create a comprehensive serials holdings display. They abandoned the use of their catalog for serials and removed all electronic journal records. The logic for this decision focused on the smaller effort required to manage print holdings in the knowledge base as opposed to the extreme effort required to catalog electronic journals for display in the online public access catalog. This solution also resolved the problem of having a different set of journals available from the catalog than the set available through the A–Z list by consolidating all serials holdings to a single point of access. Librarians at Lowe's library have received fewer holdings-related questions and have observed less confusion about where to go for serial discovery.

For many libraries, however, the adoption of a MARC record service has allowed the catalog to remain a viable option for serials access. Kemp's study investigating librarian satisfaction with MARC record services revealed that most librarians were satisfied with their implementation

of these services.[120] Librarians did complain about the number of brief records included with the service and expressed a desire for greater accuracy, but they felt that some access was better than none. Mugridge and Edmunds also discussed the benefits of batchloading MARC records sets.[121] These sets often prove valuable in revealing hidden collections and increasing collection use. After MARC record loads, librarians at Penn State would often observe an increase in collection use within days. Mugridge and Edmunds also provided a useful description of the typical workflow at Penn State for batchloading records. Similar to Kemp's observations, the authors noted record quality as a concern with these record sets. However, the desire to increase access to these collections trumps concerns about quality for Penn State. The authors stated that "batchloading allows for greater granularity—providing title-level access for collections for which only collection-level access was available previously and providing analytical access to items for which only title-level access was available."[122]

Such dramatic departures in workflows often result in unanticipated consequences, and implementation of a MARC record service is no exception. Mugridge and Edmunds as well as Kemp observed that use of a MARC record service is often incompatible with the single-record approach in serials cataloging, where multiple versions of a serial title are described using one record.[123] Libraries implementing these services have often reversed their policy from a single-record approach to a separate-record approach. Mugridge and Edwards discussed difficulties in de-duplicating records to maintain a single-record approach with a batchloading process. Considering these implementation challenges, and given the large number of electronic resources to catalog, Mugridge and Edwards argued that a single-record policy is "increasingly difficult to justify or maintain."[124] Carter discussed the dilemma of determining appropriate cataloging treatment for serials she identifies as quasi-serials; these are serials with both monographic and serial characteristics, such as standing orders and analytical series.[125] For example, batchloading MARC records for e-books often results in inconsistent cataloging treatment for these quasi-serials. Carter explained that these records are often loaded into the catalog with no check to determine whether any of the e-book titles loaded had historically received serial treatment. Given the efficiencies gained by loading MARC records sets, many libraries may need to reconsider the advantages and disadvantages of providing serial treatment for quasi-serials. To assist with this assessment, Carter provided a useful summary of these pros and cons. She commented that inconsistent bibliographic treatment of quasi-serials is just a small example of the "growing ambiguity between serials and monographs."[126]

The Influence of the Knowledge Base on Cataloging Practices

Management of a MARC record service often occurs through the same knowledge base a library uses to generate its A–Z serial list and to resolve to the appropriate copy. Curran wrote of another benefit that can be gained through a library's link resolver: using the OpenURL in the MARC record instead of using direct links.[127] Curran cited several reasons for this change: URL and coverage information are more efficiently maintained in one location, such as the knowledge base; and the OpenURL is more stable than a static URL. This practice also has allowed Curran's library to maintain a single-record approach. However, Curran admitted that occasional data errors in the knowledge base have led to bad links. She hoped the KBART initiative would lead to improved quality control within knowledge bases. Cole echoed these concerns about quality control in his review of the historical evolution of cataloging rules for minor and major title changes in comparison to publisher and aggregator treatment of title changes.[128] With publishers and aggregators often including titles with different ISSNs under the latest title, Cole observed this often can lead to access problems, especially when incorrect data are fed to a knowledge base provider. He was realistic about the vagaries of cataloging practice and does not expect publishers and aggregators to follow cataloging rules by the book. Cole did note the value of providing some level of access to both previous and current titles in knowledge base environments that result in A–Z serial lists. Page and Kuehn provided another user-focused perspective on knowledge base quality through their analysis of ILL requests.[129] In their analysis of ILL requests that were cancelled because materials were discovered locally, the authors discovered that "these requests were also associated with problematic OpenURL links to publisher or content provider Web pages."[130] All three articles indicated a need for cleaner metadata and sustained knowledge base maintenance.

Singer provided an interesting opinion piece about another possible solution to data quality concerns: the idea of a centralized knowledge base discussed in the UKSG's 2007 report "Link Resolvers and the Serials Supply Chain."[131] Singer criticized the UKSG's ideas to govern a centralized knowledge base through one organization. Instead, he believed that a community-based knowledge base would be a better option. He commented that "a centralized, standardized approach, maintained by librarians, publishers and vendors could not only reduce total cost of ownership, but also improve the quality of the data."[132] Singer pointed out numerous benefits of a community-maintained knowledge base: librarians would gain the flexibility to contribute to the knowledge base when vendors are unable or unwilling to make requested changes, it would create a single standard

framework for holdings, and it would eliminate maintenance of multiple knowledge bases in the vendor community. Ultimately, Singer emphasized that the "world does not need another closed-access, subscription based library data silo."[133]

Ensuring Perpetual Access to Electronic Content

The serials literature also focused on negotiating for and ensuring perpetual access rights to electronic resources. Articles on this topic examined the current state of perpetual access rights within license agreements, creating awareness of existing perpetual access provisions and initiatives in development to support this awareness. Rogers examined long-term access provisions in electronic journal license agreements in place for university and polytechnic libraries in New Zealand and determined that even though libraries value perpetual access and archival guarantees, these rights are often missing from negotiated license agreements.[134] Rogers found that licenses failed to address perpetual access or archival rights in 70 percent of the agreements reviewed. Zambare and colleagues also found that the provisions for these rights in license agreements were unacceptable.[135] These authors examined their license agreements for these terms after receiving content for a cancelled title in an unusable archival format. Before further attempts to go electronic-only, the library negotiated with vendors for "low or no-cost access to subscribed back files in a contemporary format."[136]

Perpetual access and archival provisions are often a source of confusion for many librarians. Keller, McAslan, and Duddy from the Oxford University Library attempted to remedy this confusion by creating a long-term access policy for their library's e-journals.[137] This document provided an overview of long-term access options for publisher-licensed content, JSTOR, OA journals, aggregators, Portico, Lots of Copies Keep Stuff Safe (LOCKSS), Controlled Lots of Copies Keep Stuff Safe (CLOCKSS), and back issues. The policy is straightforward, easy to understand, and, best of all, available to other librarians to either adopt or use as a model for a similar policy of their own. Another initiative described by Burnhill and colleagues, the Piloting an E-journal Preservation Registry Service (PEPRS), also focused on creating awareness of archival provisions for titles.[138] Sponsored by JISC, the PEPRS project is a pilot of an e-journals preservation registry through the United Kingdom academic data center and the international standards body for serials. Archiving organizations working with PEPRS include CLOCKSS, Portico, and e-Depot. This pilot is based on a study by Oppenheim and Rowland submitted to JISC in 2009.[139] Through research and interviews, this study revealed the need for a reliable information source on archiving options for journals and that librarians would most likely consult the registry while making serials management decisions to renew subscriptions, change formats, or store or withdraw print volumes. This registry would hopefully reveal "gaps in archive provision."[140] Burnhill and colleagues acknowledged concerns about funding and maintenance of the PEPRS project but, if successful, this tool could prove to be useful for the international serials community.

The Blurring and Decline of Formats

E-Books or Serials?

Serials professionals are in the business of providing metadata and acquisitions support for materials funded and issued on a continuing basis. E-books increasingly fit this description, with some purchasing models requiring subscription and annual access fees. Management of an e-book collection is not so different from managing an e-journal package; often a license agreement defines terms of use, a purchasing model must be negotiated, and a title list requires management. Thus e-book management was a common theme in the serials literature. The journal *Serials* published a special supplement on e-books in 2009 and both of the NASIG and UKSG annual conferences included sessions on e-book management.

The NASIG session "When Did (E)-Books Become Serials?" presented by Armstrong and colleagues explored how e-books are both similar and dissimilar to serials.[141] Speakers commented on the erosion of the book as a container, observing that metadata describing book content (abstracts, MARC records, and digital object identifiers) could be found at the chapter level. This is parallel to the breakdown of the journal and issue with a metadata focus on the article level. Presenters also discussed similarities of e-books to Big Deal journal packages because e-book packages can be leased, have annual fees, and allow swapping of titles in and out of the package. One speaker noted that "subscriptions are probably the most successful business model for e-books."[142] The presenters also observed that unlike serials, e-books do not yet have the infrastructure to support their sometimes continuing nature. For instance, unlike the subscription management systems offered by agents, many book vendors do not have systems in place to "control packages of books."[143] The size of e-book packages was noted as another difference between the two formats, with journals often sold in groups of hundreds and e-book packages including thousands of titles. Consequently, knowledge base management can quickly become problematic and time-consuming given the number of e-books in a package. To assist with this problem, McCracken observed that the industry "need[s] a better way of transferring content from the content providers to the electronic resources and access management services (ERAMS) vendors. Vendors

need to transmit information to knowledge base providers on behalf of libraries."[144]

The UKSG report by Thompson and Sharp also discussed the blurring lines between e-journals and e-books. They emphasized the importance of integrating e-books along with e-journals into the library's ERM tools, such as the link resolver and catalog, to increase discovery and use of these resources.[145] They also explained "students are interested in content, not format; if they have to know whether something is a journal article or a book chapter in order to search for it effectively, the potential discoverability of resources is adversely affected."[146] The blurring of content lines also was discussed by Soules in "E-Books and User Assumptions," in which she outlined several user studies to analyze user assumptions of electronic content.[147] She discussed the breakdown of terms like "e-journal" and "e-book," noting that publishers already presented mixed serial and monograph content on their platforms. Soule explained that "formats are blending; content is simply content. . . . In the growing world of information bites, let us focus on e-resource and e-content and drop terms like e-book and e-journal. It would enable us to view these individual pieces on their own merit."[148]

Print Retention and Storage

Another format that received attention in the literature is the print serial. However, unlike the articles focused on the increased use of e-books, the literature discussing print serials focused on the decline of print and issues related to retention and storage. The Ithaka report *What to Withdraw: Print Collections after Digitization* observed that print serials are no longer the format of choice for access, acknowledging that "large-scale digitization of print journal collections has led to most access needs being met via digital surrogates."[149] The role of print in today's information environment is therefore primarily one of preservation. With space at a premium at most libraries and budgets tight because of the economic recession, the Ithaka report aimed to address which print titles libraries could responsibly withdraw. Through their analysis process and interviews with librarians, the authors determined that few titles can currently be withdrawn from academic library collections. However, print versions held in two print repositories, such as those titles included in JSTOR, are candidates for withdrawal. The report discussed several reasons to retain print, including "the need to fix scanning errors; insufficient reliability of the digital provider, inadequate preservation of the digitized versions, the presence of significant quantities of important non-textual material that may be poorly represented in digital form; and campus political considerations."[150] Libraries can undertake numerous strategies to increase the number of titles they can withdraw by making the academic community aware of local preservation efforts, upgrading digitization efforts when quality is low, and participating in large-scale preservation and storage programs.

O'Connor and Jilovsky examined many of the preservation efforts that could help address concerns presented in the Ithaka report.[151] These solutions include national repositories, repository libraries, and last copy programs. One example is the Universal Repository Library (URL) vision, which grew from the International Conference on Repository Libraries and proposes to link repositories on an international scale. Another initiative discussed is ASERL's "virtual storage collection to . . . assist with the identification of last copies and the wider availability of low-use materials."[152] The authors described common threads in the literature on print storage, including the value of institutional storage facilities versus collaborative efforts, the need to analyze the costs associated with print retention, and the issue of materials ownership.[153] O'Connor and Jilovsky argued "that a network of national . . . print repositories will provide the most reliable and cost-effective solution" to print storage.[154]

The "Key Issue" column in the journal *Serials* reviewed the UK Research Reserve (UKRR) pilot project, a promising national repository initiative identified in the Ithaka Report.[155] Crawford's overview of the UKRR pilot provided a brief description of the collaboration between the British Library and higher education institutions in the United Kingdom to create a shared collection repository. Access to materials held in the print repository is provided through document delivery, and at least two additional print copies of the included titles are held by participating libraries. This project has freed more than eleven thousand meters of shelf space, allowing universities to repurpose this space for study areas, workspaces, and new collections.

Several articles in the literature discussed tools to assist deselection. Ward and Aagard described one library's process of evaluating the collection for titles to withdraw using the WorldCat Analysis tool.[156] In this example, librarians created a subject list of titles with information about duplicate holdings at peer institutions. This information was used to create criteria for deselection. Sorensen described the use of a database tool developed using Drupal, called the 5K Run Toolkit, to assist with managing a deselection project.[157] The library identified three categories of material for possible storage or withdrawal: print journals available through JSTOR were considered for disposal, print journals available online through the publisher were considered for storage, and print journals available through an aggregator also were considered for storage.

Lingle and Robinson provided a well-written, two-part case study describing a health sciences library's project to deselect print and replace high-use print journal titles

with online backfiles.[158] The demand for space to build a clinical simulation facility precipitated the transition from print to electronic-only. Part 1 of this series described how the library made their deselection decisions through usage data, an assessment of stable online content, and an overlap analysis with other university campuses. Librarians also considered the availability of online backfiles. Part 2 detailed the decision-making process using these factors, including ILL, and detailed the implementation process needed to ensure a seamless transition of the collection for patrons. The authors noted "people clearly prefer the convenience of full-text access to the journal literature from their office or remotely from anywhere with an Internet connection."[159] Users revealed, however, a continued desire for a current journal reading area and a need for more computer workstations to support access to the online collection.

Web 2.0

Influence of the Internet and Web 2.0 on Scholarly Communication

Ending this literature review as it began, with a definition from Wikipedia, seems fitting. Wikipedia defines "Web 2.0" as the "participatory Web," being associated with "web applications that facilitate interactive sharing, interoperability, user-centered design and collaboration of the World Wide Web."[160] Dodds expanded further on the fundamental concepts behind Web 2.0, such as community, collaboration, and participation, in his well-written article "The Threads of Web 2.0."[161] Dodds emphasized the value of networked services and the interweaving and integration of data across the web through the evolution of technology standards. Dodds described Web 2.0 as a "move away from the Internet as simply a platform for exchanging information, towards the Internet as a platform for creating and working with information; moving from a distribution system towards a collaborative environment."[162] He provided advice to publishers on how they can further embrace Web 2.0 technologies while providing content to users on the web. Publishers need to identify and create opportunities for connections across their user community and should allow for public sharing of information through mashup technologies to both integrate and participate within the web environment on a more global scale.

Abram also provided an overview of Web 2.0 technologies, outlining many of the applications that exemplify Web 2.0 concepts, such as RSS (really simple syndication) feeds, wikis, blogs, widgets, APIs (application programming interfaces), streamed media, and social bookmarking.[163] He included open source systems as well as the open access model for publication as examples of Web 2.0. Abram directly related Web 2.0 to librarianship through a brief discussion of the characteristics that define "Librarian 2.0," namely, the willingness to learn new tools to facilitate work, adopt and integrate the Open URL across library services, and incorporate nontraditional cataloging such as user-driven tagging and folksonomies. Abram touched on the theme of content over format, noting that librarians with Web 2.0 sensibilities should be "container and format agnostic."[164] He explained that Web 2.0 "is primarily about a much higher level of interactivity and deeper user experiences, which are enabled by the recent advances in Web software combined with insights into the transformational aspects of the Internet."[165]

The transformative effect of the web also was discussed by Kaser, who provided an example of the potential influence of Web 2.0 on electronic journal collections.[166] Kaser acknowledged the importance of the collections themselves, but noted that Web 2.0 applications "layered on top of the collection" will "get at knowledge in new and exciting ways."[167] For example, Web 2.0 applications such as social bookmarking and tagging not only enhance the discoverability of electronic journal collections, they also can personalize these collections by allowing the user to associate their personal context through descriptors and comments. Users are able to engage scholarly content to support contextual learning, a process much more valuable than simple exposure to a closed scholarly communication process. These applications provide users with a greater opportunity to become part of that scholarly community.

Examples of Web 2.0 Support For Serials Management

Numerous innovative uses of Web 2.0 applications to support management of serials and electronic resources were discussed throughout the literature. Badman and Hartman as well as Sutherland and Clark discussed the use of Web 2.0 technologies, such as RSS feeds, to create virtual journal reading rooms for patrons.[168] Badman and Hartman provided useful explanations of RSS technologies that aggregate, deliver, and organize feeds and discussed the value of creating virtual reading rooms to increase awareness of the journal collection. Both sets of authors mentioned the JISC-funded ticTOCs (a journal table of contents service) project as a great resource for aggregating journal feeds from a variety of publishers and vendors for searching and browsing.

ERM also could use Web 2.0 tools, such as blogs, twitter feeds, and chat programs, to facilitate communication between librarians and their peers as well as vendors and publishers. Emery discussed communication benefits from social networking for electronic resource librarians, including quick consultations with colleagues to resolve problems, mining for ideas to improve local workflows, and attending

virtual conferences to stay current with the field.[169] Along these same lines, Wood commented that librarians can use tools like blogs and RSS feeds to stay up-to-date about happenings and trends in the profession.[170] Numerous blogs provide information "about issues involving the acquisition and utilization of electronic resources, from licensing issues to digital rights, from copyrights to open access."[171]

Managing access to electronic resources to enhance their discoverability is another common use of Web 2.0 technologies. Kapucu, Hoeppner, and Dunlop described how the University of Central Florida used social bookmarking through Delicious to provide additional, customized access to library databases.[172] The authors discussed the value of tagging social bookmarks for "users to personalize their links, impart ad hoc organization, improve findability, and lay the ground work for social networking."[173] Churchill and colleagues described the use of social bookmarking as a core component of the Repository of Interactive Social Assets for Learning (RISAL) Project, which serves as a repository of resources such as articles, websites, and presentations used for learning and instruction.[174] The resources and bookmarks (called assets) included in the system can be tagged, embedded, linked to, ranked, and categorized to allow for sharing and reuse. Instead of these class resources living on individual student spaces or hidden within course management software, this experiment aims to create a growing network of assets and build a collaborative Web 2.0 environment to better support classroom instruction.

Other Web 2.0 examples discussed in the literature showed enhanced accessibility of electronic resources through the library catalog. Kemp discussed the integration of Web 2.0 technologies with the catalog, including mashups, tagging, and recommender features, in her discussion of advancements supporting serials cataloging.[175] Kemp described future scenarios of how these Web 2.0 applications can enhance the search experience. In another article by Singer discussing linked data, the primary focus is not necessarily library catalog functionality but the metadata within it.[176] Singer opened the column with a description of problems with library data, such as the duplication of data and the self-contained, nonrelational nature of this kind of data. He then defined the linked data "movement whose intention is to make these data structured, reusable, machine-readable, and interrelated."[177] After outlining the mechanics of linked data, such as using Uniform Resource Identifiers (URI) for naming, Singer provided a useful example of linked data within a library context, describing the potential linking effects across vended and local library systems when URIs are assigned to journal titles, database names, and organizations. He argued that libraries need to abandon the closed nature and controlled vocabularies of their data, utilizing linked data instead as a means to integrate with the web and remain relevant to the future information environment.

Conclusion

Ogburn, in "Defining and Achieving Success in the Movement to Change Scholarly Communication," discussed five stages of transition libraries need to experience to achieve a cultural shift in scholarly communication.[178] These stages include awareness, understanding, ownership, activism, and transformation. The last of these stages, transformation, "equates to attainment of a profound alteration of assumptions, methods and culture."[179] Ogburn rightly noted that this last stage is difficult to achieve, but one could argue on the basis of the serials literature of 2008 and 2009 that the serials profession is well positioned to realize and support a fundamental change in scholarly communication. The literature of this period presented the economic crisis as a catalyst for change because many librarians were experimenting with pricing models and valuing access over ownership. The economic crisis also has elevated the tensions behind the serials crisis, which almost assures serials professionals' willingness to support alternative models of scholarly communication. Other themes in the literature such as the increasing value of consortia, building of collaborative storage and national repositories, and enhanced communications through Web 2.0 technologies revealed an enhanced sense of community across the profession. Examples of information professionals embracing openness through development of open systems, support for standards, and promotion of open access also abound in the literature. The role of the web as a platform to distribute information is forcing the library and publishing communities to connect and build services on the web to stay relevant. The web has served a critical role in stripping any residual sacred cows from the serials profession, including the definition of a serial, the value of content over the container, the prioritization of access over ownership, the future of local collections, and even the future of the ever-increasing serials budget given the possibilities of an author-pays pricing model. Like it or not, serials professionals have found themselves in the midst of a transition. By embracing the core tenets of openness, such as community, interoperability, and accessibility, serialists are in fact participating in the transformation of scholarly communication.

References and Notes

1. Wikipedia, "Financial Crisis of 2007–2010," http://en.wikipedia.org/wiki/Great_Recession (accessed May 31, 2010).
2. Lee C. Van Orsdel and Kathleen Born, "Embracing Openness: Global Initiatives and Startling Successes Hint at the Profound Implications of Open Access on Journal Publishing," *Library Journal* 133, no. 7 (Apr. 15, 2008): 53–58; Lee C. Van Orsdel and Kathleen Born, "Reality Bites: In the Face of the Downturn, Libraries and Publishers Brace for Big Cuts,"

Library Journal 134, no. 7 (Apr. 15, 2009): 36–40.
3. Omnibus Appropriations Act, 2009, Public Law 111-8, 111th Cong., 1st sess. (Mar. 11, 2009), § 17; European Research Council, "ERC Scientific Council Guidelines for Open Access" (Dec. 17, 2007), http://erc.europa.eu/pdf/ScC_Guidelines_Open_Access_revised_Dec07_FINAL.pdf (accessed Sept. 24, 2010); Harvard's open access policies and numerous other university and government policies are registered with the Registry of Open Access Repository Material Archiving Policies (ROARMAP), www.eprints.org/openaccess/policysignup (accessed Sept. 24, 2010).
4. Van Orsdel and Born, "Reality Bites."
5. Van Orsdel and Born, "Embracing Openness."
6. International Coalition of Library Consortia (ICOLC), "Statement on the Global Economic Crisis and Its Impact on Consortial Licenses," July 19, 2009, www.library.yale.edu/consortia/icolc-econcrisis-0109.htm (accessed Aug. 17, 2010); Association of Research Libraries (ARL), "ARL Statement to Scholarly Publishers on the Global Economic Crisis," Feb. 19, 2009, www.arl.org/bm~doc/economic-statement-2009.pdf (accessed Aug. 17, 2010).
7. ICOLC, "Statement on the Global Economic Crisis."
8. Van Orsdel and Born, "Reality Bites," 36.
9. ARL, "ARL Statement to Scholarly Publishers."
10. Sarah Anne Murphy, "The Effects of Portfolio Purchasing on Scientific Subject Selections," *College & Research Libraries* 69, no. 4 (July 2008): 332–40.
11. William H. Walters, "Journal Prices, Book Acquisitions, and Sustainable College Library Collections," *College & Research Libraries* 69, no. 6 (Nov. 2008): 576–86.
12. Ibid., 576.
13. Ibid., 582.
14. ARL, "ARL Statement to Scholarly Publishers."
15. Jill E. Grogg and Beth Ashmore, "The Art of the Deal: Negotiating in Times of Economic Stress," *Searcher* 17, no. 6 (June 2009): 42–49.
16. Ibid., 49.
17. Ibid.
18. Bjorn Ortelbach, Sebastian Schultz, and Svenja Hagenhoff, "Journal Prices Revisited: A Regression Analysis of Prices in the Scholarly Journal Market," *Serials Review* 34, no. 3 (Sept. 2008): 190-98.
19. Joseph R. Kraus and Rachel Hansen, "Local Evaluation of Chemistry Journals," *Issues in Science & Technology Librarianship*, no. 54 (Summer 2008), www.istl.org/08-summer/article2.html (accessed Aug. 17, 2010).
20. Ibid.
21. Ibid.
22. Roger C. Schonfeld, "The Value of Scholarly Journals in a Two-Sided Market," *Serials Librarian* 54, no. 1/2 (2008): 121–26.
23. Ibid., 122.
24. Primary Research Group, *The Survey of Academic & Research Library Journal Purchasing Practices* (New York: Primary Research Group, 2008).
25. Katherine A. Perry, "Where Are Library Consortia Going? Results of a 2009 Survey," *Serials* 22, no. 2 (2009): 122–30.
26. Grogg and Ashmore, "The Art of the Deal: Negotiating."
27. Van Orsdel and Born, "Reality Bites"; Grogg and Ashmore, "The Art of the Deal: Negotiating"; "S.C. Schools May Lose Shared Databases," *American Libraries* 40, no. 3 (Mar. 2009): 20.
28. Van Orsdel and Born, "Reality Bites," 37.
29. ARL, "ARL Statement to Scholarly Publishers."
30. Jill E. Grogg and Beth Ashmore, "The Art of the Deal: The Power and Pitfalls of Consortial Negotiation," *Searcher* 17, no. 3 (Mar. 2009): 40–47.
31. Ibid., 41.
32. Tom Sanville, "Do Economic Factors Really Matter in the Assessment and Retention of Electronic Resources Licensed at the Library Consortium Level?" *Collection Management* 33, no. 1/2 (2008): 1–16.
33. See, for example, Amy Carlson and Barbara M. Pope, "'The Big Deal': A Survey of How Libraries are Responding and What the Alternatives Are," *Serials Librarian* 57, no. 4 (2009): 380–98; Zachary Rolnik, "Big Deal = Good Deal?" *Serials Librarian* 57, no. 3 (2009): 194–98; Christina Torbert, "Collaborative Journal Purchasing Today: Results of a Survey," *Serials Librarian* 55, no. 1/2 (2008): 168–83; and Cecilia Botero, Steven Carrico, and Michele R. Tennant, "Using Comparative Online Journal Usage Studies to Assess the Big Deal," *Library Resources & Technical Services* 52, no. 2 (2008): 61–68.
34. Grogg and Ashmore, "The Art of the Deal: Negotiating," 48.
35. Torbert, "Collaborative Journal Purchasing Today," 168.
36. Jill Taylor-Roe, "'To Everything There Is a Season': Reflections on the Sustainability of the 'Big Deal' in the Current Economic Climate," *Serials* 22, no. 2 (2009): 113–21.
37. Rolnik, "Big Deal = Good Deal?" 197.
38. Taylor-Roe, "'To Everything There Is a Season.'"
39. Botero, Carrico, and Tennant, "Using Comparative Online Journal Usage Studies."
40. Miquel Termens, "Looking Below the Surface: The Use of Electronic Journals by the Members of a Library Consortium," *Library Collections, Acquisitions, & Technical Services* 32, no. 2 (2008): 76–85.
41. Murphy, "The Effects of Portfolio Purchasing."
42. Ibid., 339.
43. Van Orsdel and Born, "Embracing Openness," 58.
44. Rickey D. Best, "Is the 'Big Deal' Dead?" *Serials Librarian* 57, no. 4 (2009): 353–63.
45. Colleen Cleary, "Why the 'Big Deal' Continues to Persist," *Serials Librarian* 57, no. 4 (2009): 364–79.
46. Sean O'Doherty and Bob Boissy, "Is There a Future for the Traditional Subscription-Based Journal?" *Serials Librarian* 56, no 1-4 (2009): 155–62.
47. Ibid., 161.
48. Cleary, "Why the 'Big Deal' Continues to Persist."
49. Nicole Mitchell and Elizabeth Lorbeer, "Building Relevant and Sustainable Collections," *Serials Librarian* 57, no. 4 (2009): 327–33.
50. Paul Harwood and Albert Prior, "Testing Usage-Based E-journal Pricing," *Learned Publishing* 21, no. 2 (2008): 133–39; Patrick Carr, "Acquiring Articles through Unmediated, User-Initiated Pay-Per-View Transactions: An Assessment of Current Practices," *Serials Review* 35, no. 4 (2009): 272–77.

51. Carr, "Acquiring Articles through Unmediated, User-initiated Pay-per-view Transactions," 272.
52. Harwood and Prior, "Testing Usage-Based E-journal Pricing," 133.
53. Ibid.; Clint Chamberlain and Barbara MacAlpine, "Pay-Per-View Article Access: A Viable Replacement for Subscriptions?" *Serials* 21, no. 1 (2008): 30–34; Harwood and Prior, "Testing Usage-Based E-journal Pricing."
54. Robert E. Wolverton Jr. (recorder) and Tim Bucknall (presenter), "Are Consortium 'Big Deals' Cost Effective? A Comparison and Analysis of E-journal Access Mechanisms—Workshop Report," *Serials Librarian* 55, no. 3 (2009): 469–77.
55. Harwood and Prior, "Testing Usage-based E-journal Pricing."
56. Carr, "Acquiring Articles through Unmediated, User-initiated Pay-per-view Transactions."
57. Ibid., 275.
58. Patrick L. Carr, "Forcing the Moment to Its Crisis: Thoughts on Pay-Per-View and the Perpetual Access Ideal," *Against the Grain* 21, no. 6 (Dec. 2009/Jan. 2010): 14, 16.
59. Ibid., 14.
60. Carr, "Forcing the Moment to Its Crisis."
61. Chamberlain and MacAlpine, "Pay-Per-View Article Access"; Wolverton and Bucknall, "Are Consortium 'Big Deals' Cost Effective?"
62. Harwood and Prior, "Testing Usage-based E-journal Pricing," 138.
63. Anna Hulseberg and Sarah Monson, "Strategic Planning for Electronic Resources Management: A Case Study at Gustavus Adolphus College," *Journal of Electronic Resources Librarianship* 21, no. 2 (Apr. 2009): 163–71.
64. Maria Collins, "Evolving Workflows: Knowing When to Hold'em, Knowing When to Fold'em," *Serials Librarian* 57, no. 3 (2009): 261–71.
65. Ibid., 264.
66. Kristen Blake and Erin Stalberg, "Me and My Shadow: Observation, Documentation, and Analysis of Serials and Electronic Resources Workflow," *Serials Review* 35, no. 4 (2009): 242–52.
67. C. Rockelle Strader, Alison C. Roth, and Robert W. Boissy, "E-Journal Access: A Collaborative Checklist for Libraries, Subscription Agents, and Publishers," *Serials Librarian* 55, no. 1/2 (2008): 98–116.
68. Marshall Breeding, "Helping You Buy Electronic Resource Management Systems," *Computers in Libraries* 28, no. 7 (July/Aug. 2008): 6–8, 10, 12–14, 16–18, 94, 96.
69. Maria Collins, "Electronic Journal Forum: Electronic Resource Management System (ERMS) Review," *Serials Review* 34, no. 4 (2008): 267–99.
70. Kevin Stranack, "The Researcher Software Suite: A Case Study of Library Collaboration and Open Source Software Development," *Serials Librarian* 55, no. 1/2 (2008): 117–39.
71. Ibid., 138.
72. William Doering and Galadriel Chilton, "A Locally Created ERM: How and Why We Did It," *Computers in Libraries* 28, no. 8 (Sept. 2008): 6–7, 46–48; William Doering and Galadriel Chilton, "ERMes: Open Source Simplicity for Your E-resource Management," *Computers in Libraries* 29, no. 8 (Sept. 2009): 20–24.
73. Doering and Chilton, "A Locally Created ERM: How and Why We Did It."
74. Jennifer Watson, Dalene Hawthorne (presenters), and Susan Wishnetsky (reporter), "ERM on a Shoestring: Betting on an Alternative Solution," *Serials Librarian* 54, no. 3/4 (2008): 245–52.
75. Adam Murray, "Electronic Resource Management 2.0: Using Web 2.0 Technologies as Cost-Effective Alternatives to an Electronic Resource Management System," *Journal of Electronic Resources Librarianship* 20, no. 3 (Oct. 2008): 156–68.
76. Ibid., 158.
77. Maria Collins and William T. Murray, "SEESAU: University of Georgia's Electronic Journal Verification System," *Serials Review* 35, no. 2 (2009): 80–87.
78. Taryn Resnick et al., "E-Resources: Transforming Access Services for the Digital Age," *Library Hi Tech* 26, no. 1 (2008): 141–56.
79. Kristen Blake and Jacquie Samples, "Creating Organization Name Authority within an Electronic Resources Management System," *Library Resources & Technical Services* 53, no. 2 (Apr. 2009): 94–107.
80. Marilyn White and Susan Sanders, "E-Resources Management: How We Positioned Our Organization to Implement an Electronic Resources Management System," *Journal of Electronic Resources Librarianship* 21, no. 3 (2009): 183–91; Nancy Beals, "Selecting and Implementing an ERMS at Wayne State University: A Case Study," *Journal of Electronic Resources Librarianship* 20, no. 1 (July 2008): 62–69.
81. Kristine Condic, "Uncharted Waters: ERM Implementation in a Medium-Sized Academic Library," *Internet Reference Services Quarterly* 13, no. 2 (2008): 133–45; Jill E. Grogg, "Electronic Resource Management Systems in Practice," *Journal of Electronic Resources Librarianship* 20, no. 2 (2008): 86–89; Denise Pan, "Not One-Size-Fits-All Solution: Lessons Learned from Implementing an Electronic Resources Management System in Three Days," *Journal of Electronic Resources Librarianship* 21, no. 3 (2009): 279–92; Ted Fons et al., "Real ERM Implementations: Notes from the Field," *Serials Librarian* 56, no. 1–4 (2009): 101–8.
82. Jill E. Grogg, "Electronic Resource Management Systems in Practice"; Donna F. Ekart, "Somewhere over the Verde Rainbow," *Computers in Libraries* 28, no. 2 (Sept. 2008): 8, 10, 44–45.
83. Ekart, "Somewhere Over the Verde Rainbow," 45.
84. Pan, "Not One-Size-Fits-All Solution."
85. Pan, "Not One-Size-Fits-All Solution"; Fons et al., "Real ERM Implementations"; Condic, "Uncharted Waters"; White and Sanders, "E-Resources Management."
86. Bonnie Tijerina and Douglas King, "What is the Future of Electronic Resource Management Systems?" *Journal of Electronic Resources Librarianship* 20, no. 3 (Oct. 2008): 147–55.
87. Ibid., 153.
88. Oliver Pesch, "ISSN-L: A New Standard Means Better Links," *Serials Librarian* 57, no. 1/2 (2009): 40–47.
89. Ibid., 47.

90. Sophie Vincent, "The New Edition of the ISSN International Standard Makes Life Easier for the Serials Community," *Serials* 21, no. 1 (2008): 45–48.
91. Peter McCracken (presenter) and Michael A. Arthur (recorder), "KBART: Best Practices in Knowledgebase Data Transfer," *Serials Librarian* 56, no. 1–4 (2009): 231.
92. Norm Medeiros et al., *White Paper on Interoperability between Acquisitions Modules of Integrated Library Systems and Electronic Resource Management Systems* (Washington, D.C.: Digital Library Federation, 2008), www.diglib.org/standards/ERMI_Interop_Report_20080108.pdf (accessed Apr. 15, 2010).
93. Ibid., 4.
94. Ed Riding, "Cost of Resource Exchange (CORE): The Making of a Library Standard," *Serials Librarian* 57, no. 1/2 (2009): 51–56.
95. Ibid., 56.
96. Oliver Pesch, "An Update on COUNTER and SUSHI," *Serials Librarian* 55, no. 3 (2008): 366–72.
97. Tansy Matthews, "Where Does It Come From? Generating and Collecting Online Usage Data: An Update," *Serials* 22, no. 1 (2009): 49–52.
98. Richard Gedye, "Measuring the Usage of Individual Research Titles," *Serials* 22, no. 1 (2009): 24–27.
99. National Information Standards Organization (NISO), *SERU: A Shared Electronic Resource Understanding* (Baltimore: NISO, 2008), www.niso.org/publications/rp/RP-7-2008.pdf (accessed Aug. 17, 2010).
100. Ibid., 1.
101. Clinton K. Chamberlain, "Breaking the Bottleneck: Using SERU to Facilitate the Acquisition of Electronic Resources," *College & Research Library News* 69, no. 11 (Dec. 2008): 700–702; Zachary Rolnik and Selden Lamoureux (presenters) and Kelly A. Smith (recorder), "Alternatives to Licensing of E-Resources," *Serials Librarian* 54, no. 3/4 (2008): 281–87.
102. NISO, *SERU*; Chamberlain, "Breaking the Bottleneck."
103. Ed Pentz and Louise Cole, "The UKSG TRANSFER Project: Collaboration to Improve Access to Content," *Serials* 22, no. 2 (2009): 161–65.
104. Sheila S. Intner with Peggy Johnson, *Fundamentals of Technical Services Management* (Chicago: ALA, 2008).
105. Peter M. Webster, *Managing Electronic Resources: New and Changing Roles for Libraries* (Oxford, U.K.: Chandos, 2008).
106. Maria D. D. Collins and Patrick L. Carr, eds., *Managing the Transition from Print to Electronic Journals and Resources: A Guide for Library and Information Professionals* (New York: Routledge, 2008).
107. Becky Albitz, *Licensing and Managing Electronic Resources* (Oxford, U.K.: Chandos, 2008).
108. Holly Yu and Scott Breivold, eds., *Electronic Resource Management in Libraries: Research and Practice* (Hershey, Penn.: IGI Global, 2008).
109. Lori J. Terrill, "A Snapshot of Early Acceptance of the CONSER Standard Record in Local Catalogs," *Serials Review* 35, no. 1 (2009): 26.
110. Ibid., 16.
111. Rebecca Kemp, "Catalog/Cataloging Changes and Web 2.0 Functionality: New Directions for Serials," *Serials Librarian* 53, no. 4 (2008): 91–112.
112. Ibid., 98.
113. Ibid.
114. Mary Curran, "E-Ventures: Notes and Reflections from the E-Serials Field. Serials in RDA: A Starter's Tour and Kit," *Serials Librarian* 57, no. 4 (2009): 306–23.
115. Kemp, "Catalog/Cataloging Changes and Web 2.0 Functionality."
116. Xiaothan Chen and Stephen Wynn, "E-Journal Cataloging in an Age of Alternatives: A Survey of Academic Libraries," *Serials Librarian* 57, no. 1/2 (2009): 109.
117. Ibid.
118. Ibid., 110.
119. M. Sara Lowe, "Clear as Glass: A Combined List of Print and Electronic Journal in the Knowledge Base," *Journal of Electronic Resources Librarianship* 20, no. 3 (Oct. 2008): 169–77.
120. Rebecca Kemp, "MARC Record Services: A Comparative Study of Library Practices and Perceptions," *Serials Librarian* 55, no. 3 (2008): 379-410.
121. Rebecca L. Mugridge and Jeff Edmunds, "Using Batchloading to Improve Access to Electronic and Microform Collections," *Library Resources & Technical Services* 53, no. 1 (Jan. 2009): 53–61.
122. Ibid., 61.
123. Mugridge and Edmunds, "Using Batchloading to Improve Access to Electronic and Microform Collections"; Kemp, "MARC Record Services."
124. Mugridge and Edmunds, "Using Batchloading to Improve Access to Electronic and Microform Collections," 55.
125. Kathy Carter, "Serial or Monograph: Who Decides?" *Serials Librarian* 57, no. 1/2 (2009): 87–95.
126. Ibid., 94.
127. Mary Curran, "Single and Open: Using the SFX A–Z list URL in the Catalog," *Serials Librarian* 55, no. 3 (2008): 351–65.
128. Jim Cole, "From the Majors to the Minors: Playing the Game as the Rules are Changed," *Serials Librarian* 55, no. 1/2 (2008): 45–58.
129. Jessica R. Page and Jennifer Kuehn, "Interlibrary Service Requests for Locally and Electronically Available Items: Patterns of Use, Users, and Cancelled Requests," *portal: Libraries and the Academy* 9, no. 4 (Oct. 2009): 475–89.
130. Ibid., 475.
131. Ross Singer, "The Knowledgebase Kibbutz," *Journal of Electronic Resources Librarianship* 20, no. 2 (Sept. 2008): 81–85; see also UKSG, "Link Resolves and The Serials Supply Chain: Final Report for UKSG," May 21, 2007, www.uksg.org/projects/linkfinal (accessed Aug. 17, 2010).
132. Ibid., 81.
133. Ibid., 85.
134. Sam Rogers, "Survey and Analysis of Electronic Journal Licenses for Long-Term Access Provisions in Tertiary New Zealand Academic Libraries," *Serials Review* 35, no. 1 (2009): 3–15.
135. Aparna Zambare et al., "Assuring Access: One Library's Journey from Print to Electronic Only Subscriptions," *Serials Review* 35, no. 2 (2009): 70–74.
136. Ibid., 71.

137. Alice Keller, Jonathan McAslan, and Claire Duddy, "Long-Term Access to E-Journals: What Exactly Can We Promise Our Readers?" *Serials* 21, no. 1 (2008): 49.
138. Peter Burnhill et al., "Piloting an E-Journals Preservation Registry Service (PEPRS)," *Serials* 22, no. 1 (2009): 53–59.
139. Charles Oppenheim and Fytton Rowland, *Scoping Study on Issues Relating to Quality-Control Measures within the Scholarly Communication Process* (Bristol, UK: JISC, 2009), www.jisc.ac.uk/whatwedo/topics/opentechnologies/openaccess/reports/qualitycontrol.aspx (accessed May 2, 2010).
140. Ibid., 1.
141. Kim Armstrong et al., "When Did (E)-Books Become Serials?" *Serials Librarian* 56, no. 1–4 (2009): 129–38.
142. Ibid., 130.
143. Ibid., 134.
144. Ibid., 135.
145. Sarah Thompson and Steve Sharp, "E-Books in Academic Libraries: Lessons Learned and New Challenges," *Serials* 22, no. 2 (2009): 136–40.
146. Ibid., 139.
147. Aline Soules, "E-Books and User Assumptions," *Serials* 22, no. 3, suppl. 1 (2009): S1–S5.
148. Ibid., S4.
149. Roger C. Schonfeld and Ross Housewright, *What to Withdraw? Print Collections Management in the Wake of Digitization* (New York: Ithaka, 2009), www.ithaka.org/ithaka-s-r/research/what-to-withdraw/What to Withdraw - Print Collections Management in the Wake of Digitization.pdf, 2 (accessed Aug. 17, 2010).
150. Ibid., 2.
151. Steve O'Connor and Cathie Jilovsky, "Approaches to the Storage of Low Use and Last Copy Research Materials," *Library Collections, Acquisitions & Technical Services* 32, no. 3–4 (2008): 121–26.
152. Ibid., 123.
153. Ibid.
154. Ibid., 121.
155. Jean Crawford, "Securing Access to Print: The UKRR," *Serials* 21, no. 3 (2008): 232–34.
156. Suzanne M. Ward and Mary C. Aagard, "The Dark Side of Collection Management: Deselecting Serials from a Research Library's Storage Facility Using WorldCat Collection Analysis," *Collection Management* 33, no. 4 (2008): 272–87.
157. Charlene Sorensen, "The 5K Run Toolkit: A Quick, Painless, and Thoughtful Approach to Managing Print Journal Backruns," *Serials Review* 35, no. 4 (2009): 228–34.
158. Virginia A. Lingle and Cynthia K. Robinson, "Conversion of an Academic Health Sciences Library to a Near-Total Electronic Library: Part 1," *Journal of Electronic Resources in Medical Libraries* 6, no. 3 (2009): 193–210; Virginia A. Lingle and Cynthia K. Robinson, "Conversion of an Academic Health Sciences Library to a Near-Total Electronic Library: Part 2," *Journal of Electronic Resources in Medical Libraries* 6, no. 4 (2009): 279–93.
159. Lingle and Robinson, "Conversion of an Academic Health Sciences Library to a Near-Total Electronic Library: Part 2," 289.
160. Wikipedia, "Web 2.0," http://en.wikipedia.org/wiki/Web_2.0 (accessed July 9, 2010).
161. Leigh Dodds, "The Threads of Web 2.0," *Serials* 21, no. 1 (2008): 4–8.
162. Ibid., 5.
163. Stephen Abram, "Social Libraries: The Librarian 2.0 Phenomenon," *Library Resources & Technical Services* 52, no. 2 (April 2008): 19–22.
164. Ibid., 21.
165. Ibid., 20.
166. Dick Kaser, "Focusing on E-journals (and Blogging at the Same Time)," *Computers in Libraries* 29, no. 2 (Feb. 2009): 33–35.
167. Ibid., 35.
168. Derik A. Badman and Lianne Hartman, "Developing Current Awareness Services: Virtual Reading Rooms and Online Routing," *College & Research Libraries News* 69, no. 11 (Dec. 2008): 670–72; Michael Sutherland and Jason Clark, "Virtual Journal Room: MSU Libraries Table of Contents Service," *Computers in Libraries* 29, no. 3 (Feb. 2009): 6–7, 42–43.
169. Jill Emery, "All We Can Do Is Chat Chat: Social Networking for the Electronic Resources Librarian," *Journal of Electronic Resources Librarianship* 20, no. 4 (2008): 205–9.
170. M. Sandra Wood, "Using Blogs to Stay Current about Electronic Resources," *Journal of Electronic Resources in Medical Libraries* 6, no. 3 (2009): 242–44.
171. Ibid., 242.
172. Aysegul Kapucu, Athena Hoeppner, and Doug Dunlop, "Getting Users to Library Resources: A Delicious Alternative," *Journal of Electronic Resources Librarianship* 20, no. 4 (Dec. 2008): 228–42.
173. Ibid., 229.
174. Daniel Churchill et al., "Social Bookmarking-Repository-Networking: Possibilities for Support of Teaching and Learning in Higher Education," *Serials Review* 35, no. 3 (2009): 142–48.
175. Kemp, "Catalog/Cataloging Changes and Web 2.0 Functionality."
176. Ross Singer, "Linked Library Data Now!" *Journal of Electronic Resources Librarianship* 21, no. 2 (2009): 114–26.
177. Ibid., 117
178. Joyce L. Ogburn, "Defining and Achieving Success in the Movement to Change Scholarly Communication," *Library Resources & Technical Services* 52, no. 2 (April 2008): 44–53.
179. Ibid., 45.

Seeing Versus Saving

Recommendations for Calculating Research Use-Lighting for Library Special Collections

Lauren Anne Horelick, Ellen Pearlstein, and Holly Rose Larson

The research presented in this paper describes the measurement of light and ultraviolet energy within a special collections facility, with the goal of evaluating whether levels recommended for museums and archival collections are being exceeded during research usage. An Elsec 764 hand-held light meter was used to record the light intensity falling on collection material held within and without V-shaped book mounts and with sequential lights turned on, as occurs in collections' use. The authors developed a simple algebraic formula to calculate cumulative doses of light and incident ultraviolet radiation to determine how many hours collection material could be accessed and illuminated before damage could be expected. The authors calculated the maximum cumulative doses possible based on numbers of access hours and compared these to recommended doses for sensitive media as a monitoring strategy for the long-term preservation of light sensitive special collection materials. The results from this study suggest that the light levels evaluated are not in excess of recommended values and that the use of book mounts reduces the amount of light falling on collection material. Monitoring actions are recommended for institutions wishing to replicate the study.

Lauren Anne Horelick (l.horelick@gmail.com) is the Mellon Fellow in objects conservation at the Smithsonian's National Museum of the American Indian, Washington, D.C.; **Ellen Pearlstein** (epearl@ucla.edu) is Associate Professor in Information Studies, University of California Los Angeles and the UCLA/Getty Conservation Program, Los Angeles; and **Holly Rose Larson** (holly rose13@gmail.com) is Processing Archivist, the Autry National Center of the American West, Los Angeles.

Submitted August 27, 2010; returned to authors August 16, 2010, with request to revise; revision submitted November 12, 2010, and accepted for publication.

The authors gratefully acknowledge the work of Stefan Michalski at the Canadian Conservation Institute (CCI). The title and content of this research paper are directly influenced and rely heavily on his 2010 article "Light, Ultraviolet and Infrared," published through CCI.

The potential loss of value to library special collections materials damaged by inappropriate lighting can be estimated with simple measures described in this paper. This research aims to fill a gap in the literature concerning light damage to diverse library collections, with a specific goal of understanding the relationship between variable illumination falling on materials during research use and predicting fading behavior. Additionally, the authors explored the applicability of book mounts to determine whether their use reduces the amount of light falling on surfaces, thereby permitting greater long-term preservation of collection material. Despite numerous published studies recommending specific ranges for library and special collections storage and display illumination, no comprehensive assessment of the cumulative effects of light falling on collections accessed during use and under varying lighting conditions has been conducted.

The authors set out to measure the maximum possible annual light and ultraviolet radiation doses experienced by special collections materials through normal use by undertaking a study at the Young Research Library at the University of California, Los Angeles (UCLA), in a special collections reading room. The findings from the study stress the need for library stewards to know the duration and intensity of light falling on their most sensitive collections because this knowledge

will ultimately affect the condition of certain materials and inform care practices. Potential responses to managing lighting, such as instituting the use of lighting logs for individual items or the use of lower wattage bulbs, neutral density filters, and ultraviolet filters, are explored elsewhere in the museum lighting literature and are therefore omitted here.[1] The method used, particularly for evaluating lighting where materials are accessed, can be applied to other special collections libraries for both planning and comparative purposes.

This study was inspired by the paradox that librarians encounter—concern with providing access to information coupled with needing to restrict this access in the interest of long-term preservation. To set the context, the authors introduce how light is measured, its sources, and the importance of its cumulative and irreversible effects on special collections materials.

Literature Review: Evolution of Library Lighting Recommendations

The authors examined the literature of environmental preservation guidelines for libraries, archives, historic homes, and museums to chronicle trends in lighting recommendations. Two trends are encountered in the literature reviewed. The first is the adoption of broad ranges of light levels recommended for library reading rooms and stacks with slightly more restrictive light levels for task lighting. The second reveals guidelines and recommendations written specifically about exhibitions for libraries, including those put forward by national libraries and conservation and lighting organizations. The recommendations of lighting organizations more closely align with museum guidelines.

The idea of what is needed to implement best practices in environmental preservation for libraries, archives, and museums has changed during the past sixty years, with numerous and contradictory recommendations put forth.[2] Before detailing these recommendations, the authors first explore the terminology used throughout the literature. Beginning in the 1950s, color scientists established that visual comfort requires a minimum of 50 lux, which is enough light to ensure that the human eye is operating with the full range of color vision.[3] "Lux" is the term used to describe visible light expressed as lumens (light) per square meter.[4] Ultraviolet light is a proportion of visible energy expressed as microwatts per lumen (μW/lm).[5]

Sources agree that lux levels should be set to ensure the most effective balance between the needs of readers and the need to minimize light damage to collections.[6] Light damage is determined by the wavelength of the light, length of the exposure, and the intensity of the illumination, which in excess can accelerate deterioration on the molecular level, resulting in cumulative and irreversible damage.[7] From 1998 to the present, a range of recommendations concerning light levels for libraries and archives have been put forward by authors and private organizations. A guide to library environmental monitoring and control published in 1998 provides specifications of 200–300 lux as an acceptable range for library reading rooms and a more conservative estimation of 50–70 lux for a maximum of 60–90 days for light-sensitive materials.[8] In 2004, Ogden (in a manual for collections preservation in library design) advised values below 300–650 lux for task lighting and between 65–375 lux for inactive lighting of stacks.[9] Despite the range in lux provided, Ogden recommended that "light levels for the stacks should be set to the minimum acceptable to enable book titles and call numbers to be read."[10] A library design guideline published in 2005 with recommendations from the Illuminating Engineering Society distinguished between light levels of 500 lux for active task lighting to permit visibility of fine printed material, 50 lux for storage spaces, and 300 lux for usage spaces depending on the distance from the measured light source.[11] Beech advocated no more than 50 lux for the exhibition of collection material in a paper given at the 2005 Philatelic Congress of Great Britain.[12] A small pamphlet produced in 2007 by Etheredge recommended that libraries "put your collections in a room without windows" to ensure long-term preservation of library special collections.[13] These recommendations illustrate the ranges encountered in the literature, in which the suggested lux levels can range from as high as 650 lux to no light at all for both the use and display of library collections. Despite the numerous suggestions for lux levels in the literature, advice concerning the complex issues that arise with illuminating special collection material accessed during research use is lacking.

Much of the literature for museums and library special collections offers guidelines and recommendations specifically related to the exhibition of unique artifacts. This is largely because both museum collections and special collections spend the majority of their time in dark storage, protected from both light and ultraviolet energy. However, unlike museum objects, special collections receive their greatest light exposure during research use rather than display, and library preservation specialists note that light levels should differ within storage, exhibition, and reading rooms.[14] National standards for the display of library and archival materials approved in 2001 by the American National Standard /National Information Standards Organization (ANSI/NISO) state that for exhibitions most items should be exposed to 50–100 lux, with cumulative annual light exposures of between 42,000–84,000 lux-hours over a twelve-week period, at ten hours a day.[15] A twelve-week exhibition at these times and intensities is recommended only once every two years

so that the 42,000–84,000 lux-hours are reduced by half (21,000–42,000) when expressed as an annual total dose. These standards also suggest that at the 50–100 lux level, ultraviolet light exposure should not exceed 75µW/lumen, although Saunders recommends 10µW/lumen.[16] These standards are understandably stringent, as in the case of highly sensitive collections illuminated at 50 lux; this exhibition duration would restrict exposure to 420 hours rather than 3,000 hours of annual display.

Museum conservators in the early 1990s developed lighting policies for the display of collections with annual recommendations of total dosages, which are calculated on the basis of the sensitivity of different media.[17] Differing sensitivities are recognized for library and archival materials so that annual cumulative dosages of light have been adapted for their display.[18] Environmental guidelines for the exhibition of library and archival materials mandated in a directive to the National Archives and Records Administration include restricting the cumulative dose of light and ultraviolet radiation, while normal office lighting modified to exclude all ultraviolet radiation is recommended for research rooms in archival facilities.[19] Preservation guidelines described for special collections storage include the elimination of all natural light and the use of ultraviolet filters on all fixtures.[20] An example of a previous study applying these principles is an environmental assessment of the rare book collection at the University of Tennessee, in which the authors identified collections receiving ultraviolet radiation in excess of recommended values.[21]

Despite past and recent recommendations of environmental standards for the storage and display of library and special collections, the authors of the current study found no published strategies for determining the lighting dose accumulated by special collections during research use. This is a necessary step if cumulative doses are to be compared with recommended annual exposures, thereby permitting an evaluation of damage induced by light.

Special Collections: Issues of Restrictions and Access

Special collections in institutions are created using individually developed collection policies, with preservation of the artifact playing an overarching role.[22] Factors such as monetary value, age, physical characteristics, bibliographic value, and condition are considered by individual institutions to determine whether the item requires the preservation and security handling of a special collections department.[23] Most special collections libraries require users to sign an agreement that includes restrictions on food and beverages, requirements about approved note-taking materials (pencils instead of pens), the use of book mounts, and restrictions on photocopying. UCLA's Charles E. Young Research Library Department of Special Collections requires individual users to complete a registration form with their contact information, proof of identification, and statement of purpose.[24] Additional protocols include specifying that only special collections staff may perform duplication of material, and written protocols about reproduction state that "limits have been established to ensure preservation of the materials."[25]

Restrictions on access and use are necessary in special collections libraries because, unlike general library materials that circulate, special collections are often the most highly valued because of their rarity and preciousness. Unique items may enter the special collections division of the library immediately after acquisition, or they may enter circulating collections and be moved later to a special collection. Materials at the UCLA Department of Special Collections at the Charles E. Young Research Library, where the research was conducted, include illuminated and palm manuscripts, scrapbooks, artists' books, early newspapers, a variety of photographic media, original art works, sheet music, architectural renderings, weapons, cuneiform tablets, historic clothing, and wire recordings. Some of these materials include light-sensitive pigments on organic media, books and graphic documents, albumen prints, color photographs, parchment, leather, and new media.

The same use, display, and storage guidelines suggested for museum materials may be advisable for special collections materials because these items represent the sections within the library in which the materials parallel those found in museums. Despite the irreparable damage done by excesses of light, illumination is a necessary factor for viewing collection materials and must, therefore, be monitored and controlled.

Seeing Versus Saving

In the fall of 2008, the authors conducted a light monitoring study in the Ahmanson Murphy Reading Room at the UCLA's Charles E. Young Research Library. The study involved taking readings of visible light and ultraviolet energy with an Elsec 764 hand-held light meter, with book collections supported both inside of V-shaped book mounts and lying flat. The authors calculated the cumulative exposure of light energy to quantify the time until a just-noticeable fade (JNF) might be detected in the library's collection material. They sought to determine a strategy for measuring the lighting dose accumulated by special collections during research use to permit collections staff to compare cumulative lighting dosages to sensitivity standards. Before launching into a discussion of how the measurement and interpretation of light values are applied in this case study, defining terms used to describe light energy is necessary.

Light Energy and Ultraviolet Radiation: Definitions and Effects

The scientific and descriptive terminology used to describe light energy and its effects on special collection materials help to identify and define problems with inappropriate light levels. Light energy within the visible portion of the spectrum (figure 1) is measured between 400 and 700 nanometers (nm) and is well established as damaging to collection materials.[26] Ultraviolet radiation is measured between 200 and 400 nm, is often more damaging than visible light, and is proportional to visible energy.[27] Ultraviolet radiation is measured as the ratio of ultraviolet radiation intensity (in radiometric units $\mu W/m^2$) to light intensity (in photometric units, lux = lumen/m^2), hence the result is expressed as microwatts per lumen ($\mu W/lm$).[28] The light energy from visible and ultraviolet regions can be absorbed by the molecules within an object, causing many possible sequences of chemical reactions, known as photochemical deterioration, which is very damaging to paper.[29]

Damage to library materials from inappropriate lighting will result in irreversible condition issues that ultimately can affect intrinsic and research value. Visible light radiation originates from sunlight and some fluorescent and incandescent bulbs; the latter can cause warming and desiccation of objects if used in close proximity to collections.[30] Infrared radiant energy is seldom sufficient to induce the chemical reactions that are normally encountered in photochemical deterioration, but it can raise temperatures significantly, which can cause a different kind of damage.[31] Garrison and Lull reported that light damage and direct light on collections causes an increase in temperature, fading, yellowing, embrittlement, and weakness.[32] Saunders reported that light damage results in loss of color and strength.[33]

Measuring Light: Interpretation and Use of Values

Light is measured in museum and library contexts with devices called light meters, which contain a photoelectric cell of the type found in cameras.[34] Light meters can be used to record incidental lighting or can be set up to record lighting over a set amount of time with devices called data loggers. Light intensity is measured as visible energy per unit area, expressed as lumens per square foot (footcandle) or, using metric measurements, lumens per square meter (lux). Cumulative exposure is measured as intensity times duration, equaling lux hours (or footcandle hours), with larger quantities expressed as kilolux (Klux) hours or megalux (Mlux) hours.[35] A shift within museums and historic houses from conservation monitoring of incident light and ultraviolet levels to logging cumulative doses occurred in the late 1980s in recognition of how damage actually occurs.[36] Hain noted in 2003 that environmental monitoring of special collections had been augmented by data loggers, which collect light data alongside the temperature and relative humidity data more common to library environmental management.[37] Morris restated the advantages of logging in 2009.[38]

Specialists in museum lighting have proposed various recommended cumulative light exposures for collections.[39] Cumulative light exposure is expressed in terms of a total annual exposure, which is calculated on the basis of typical museum exhibition periods of twelve weeks with an exposure of ten hours per day. Using this type of calculation, Michalski provided the numbers of lux hours (or light intensity multiplied by time) before a JNF would occur for items of differing International Organization for Standardization (ISO) sensitivities (table 1).[40] Material sensitivities are expressed in terms of ISO Blue Wool Standards, using ISO groupings of material sensitivities. ISO Blue Wool Standards

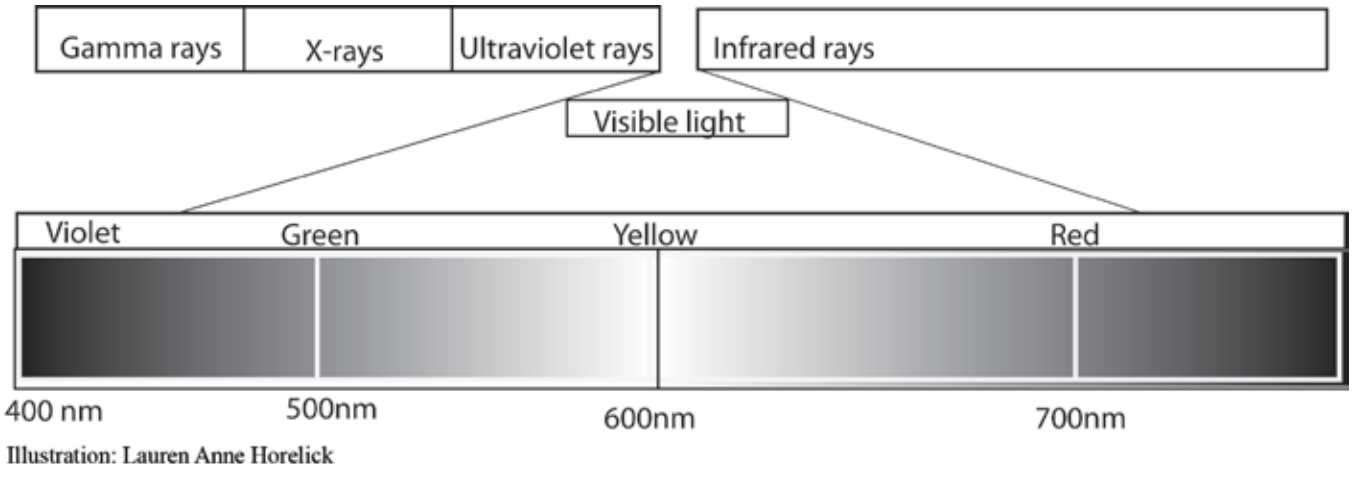

Figure 1. Electromagnetic Spectrum

Table 1. Categories of ISO Sensitive Materials and Associated Fade Time

Category	Examples of Artifacts	Estimated Years to Just Noticeable Fade
High sensitivity ISO 1,2,3	Albumen prints color photographs, plant dyestuffs, 20th c. dyes, tinted papers, ballpoint ink	1.5 to 20 years 0.23 - 3 Mlux hours 1.5 years at 150,000 lux hours per year= 225,000 lux hours till fade
Moderate sensitivity ISO 4,5,6	Parchment, leather, fur, feathers, textiles	20-700 years 3 - 105 Mlux hours 20 years at 150,000 lux hours per year = 3 million lux hours till fade
Low sensitivity ISO 7,8 above	Stone, metals, ceramics, B/W photographs, graphite, mineral pigments	300-7000 years 45 - 1050 Mlux hours 300 years at 150,000 lux hours per year 45 million lux hours till fade

Source: Stefan Michalski, Ten Agents of Deterioration: Light, Ultraviolet and Infrared: How Much Light Do We Need to See? (Ottawa: Canadian Conservation Institute, 2010), table 3, www.cci-icc.gc.ca/crc/articles/mcpm/chap08-eng.aspx#Light (accessed Oct. 9, 2010).

are dyed textiles prepared in a series of eight, with ISO 1 being the most sensitive and ISO 8 being the least sensitive.[41] Fading of these standards corresponds to generally predictable exposure levels of combined light and ultraviolet radiation, allowing users to estimate the cumulative dose that caused particular dyes in the series to fade.[42]

The reciprocity principle, which describes photochemical damage as a factor of both lighting intensity and duration, implies that damage can be mitigated by controlling one or both of these factors.[43] The reciprocity principle may be used to manipulate the variables of visible light and time to achieve necessary illumination while not exceeding cumulative annual light exposures. This concept states that limited exposure to a high-intensity light will produce the same amount of damage as a long exposure to a low-intensity light.[44] An example would be that exposing an illuminated manuscript to 100 lux for five hours would cause the same amount of damage as if it were exposed to 50 lux for ten hours. Adequate lighting is necessary to facilitate viewing, but excessive or inappropriate illumination has an irreversible effect on collection materials. While lighting standards are readily implemented in the case of museum exhibitions, applying lighting standards to special collection reading rooms in libraries is difficult because of the unique needs of researchers.

Case Study: Measuring Variable Illumination of Special Collections Materials with and without Book Mounts

The authors evaluated whether specific lighting standards or protocols are necessary in a special collections reading room by recording measurements of visible and ultraviolet light in the Charles C. Young Research Library Ahmanson Murphy Reading Room on the UCLA campus. See figures 2 and 3 for a floor plan of the reading room. This reading room houses a collection of fifteenth-century Italian manuscripts in wooden cabinets lining the walls of the room and is the sole location where users access the diverse range of special collections materials. The reading room is on the lower level of the Young Research Library and receives no direct or indirect sun light. The collection materials are stored below the reading room in the library and are brought up to the reading room when paged for use. Paged materials are held in the back of the room behind a permanently installed folding screen and are stored on wooden or metal shelves for up to five days, unless special arrangements are made by individual patrons for longer use. The materials are often housed in opaque protective enclosures, stored either vertically or horizontally.

The reading room is illuminated by four different light sources (see figure 2). Forty-two-watt compact fluorescent bulbs are located in the twelve recessed lights situated at the north and south ends of the room—six over the student monitor desk and another six over the storage area behind the screen. Three pendant lamps illuminate the main table, and each also use two 42-watt compact fluorescent bulbs. The reading room has six large conjoined wooden tables capable of accommodating twelve individuals, each with a user-operated desk lamp. Each user-operated desk lamp has one 15-watt incandescent bulb. A row of more than 200 5-watt incandescent bulbs lines the top periphery of the room's bookcases. The recessed lights, pendant lamps, and perimeter lights are used Monday through Friday for ten hours a day.

The authors took light measurements with the main reading room table's sequential user-operated desk lamps turned on to approximate the lux generated when the room is fully occupied. Light readings were taken with and without the use of foam V-shaped book mounts to determine whether the angle of illumination in normal use, which includes book mounts, increased or diffused

the total lux hitting collection material. This set of light-monitoring readings was developed to evaluate the risk of overexposure for extremely light-sensitive materials. The potential for overexposure necessitates changes in policy, such as use in a specially lit location, usage logs, or further use restrictions.

Research Method

An Elsec 764 light meter (see table 2 for specifications) was used to gather incidental, quantitative measurements of the amount of visible light and ultraviolet radiation falling on collection materials brought onto the main table in the reading room. In developing the light-monitoring method, the authors considered the potential for increased exposure from individually operated nearby light sources. They took readings from one location on the main table with sequential and adjacent lights turned on. This was done to determine whether these lights, in addition to overhead and ambient lights, increased the total lux hitting collection materials (figure 3 indicates where the readings were taken). Measurements recorded lighting incident on collections during research use and do not include exposure calculations for Italian manuscripts stored in glass front cabinets in the reading room. The authors recorded all visible and ultraviolet energy measurements while laying the Elsec 764 flat on the table, simulating the position of collection items, and then took a second round of readings in the same location measuring the amount of light within foam V-shaped book mounts.

The measured lux, both within and without book mounts, and with sequential illumination is presented in table 3. The authors used the recorded lux from these tables to determine how many hours until a JNF would occur at the given light levels by creating a simple algebraic formula:

$E/L = T_M$

Where,

E = Estimated cumulative lux hours to a just noticeable fade (based on Michalski's ISO 1–8 recommendations. See table 1 for values).

L = Measured incident lux (in a given lighting scenario).

T_M = Total hours (before a JNF occurs).

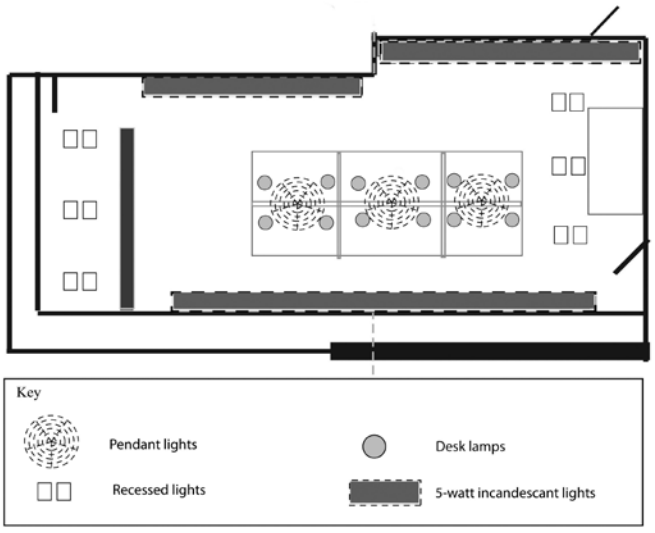

Figure 2. Ahmanson Murphy Reading Room Floor Plan Showing Light Sources

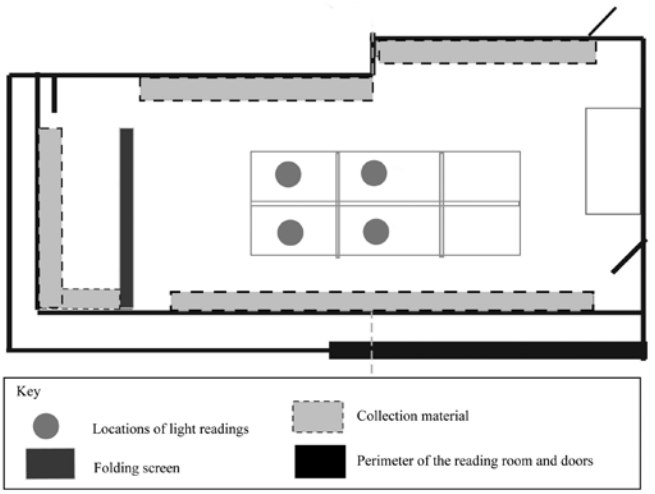

Figure 3. Ahmanson Murphy Reading Room Floor Plan Showing Locations of Light Readings

Table 2. Specifications of the Elsec764

Device	Visible Wavelength Range	Visible Power Range	UV Wavelength Range	Operating Temperature	Accuracy
Elsec 764C	400–700nm (CIE response). No correction required for different light sources.	0.1–200,000 Lux (0.1–20,000 Foot-candles).	300–400 nm	0–50°C	*Light*: 5% +1 displayed digit *UV*: 15% *Temperature*: 0.5oC (0.9oF) *RH*: 3.5%

Source: Pacific Data Systems, ELSEC 764 and 764C, www.pacdatasys.com.au/dataloggers/ELSEC_Specs.htm (accessed Nov. 13, 2010).

To calculate the hours until a JNF occurs, the authors used Michalski's annual dose of 150,000 lux hours (i.e., 50 lux at 3,000 hours) as a starting point and his approximate evaluations of years to fade for ISO sensitive materials (see table 1).[45] The values to JNF were calculated using the above formula and are shown in table 4. The cumulative lux hours predicted to cause JNF over the lifetime of different classes of ISO materials with varying light sensitivities were divided by the cumulative lux measured for routine illumination in the special collections reading room. The value generated from this formula is the number of hours until JNF occurs. The total hour maxima (T_M) should fall under the recommended ISO luminous exposures for the particular sensitivity level of each item to achieve a JNF.

To use this formula, one would first take an incident light reading, the number of which would be used in the "L" position of the formula. For example, the ambient light in the Young Research Library was recorded at 159 lux. To determine the value for "E," numbers are substituted from the ISO sensitivities (high, low, and moderate). For materials that fall under the high sensitivity, the value for "E" is 225,000. For ISO moderate-sensitivity materials, "E" is substituted with 3 million, and for low-sensitivity materials, "E" represents 45 million. To determine the JNF of highly sensitive materials, the equation would read 225,000/159 = T_M. Once solved, T_m (total hours before JNF) is 1,415 hours.

Findings and Discussion

The results from the case study to determine the effects of various illuminations falling on collection material yielded two data sets: lux values recordings and recordings of hours till JNF occurs with and without the use of V-shaped book mounts. The findings established concrete numbers representing hours until fade will occur, thereby allowing library stewards to establish reasonable numbers of usage hours per year for their collections.

Table 3 summarizes the results of the incident readings measured on the main reading room table with and without V-shaped book mounts and with successive user operated lights turned on in addition to ambient room illumination. The top portion of table 3 shows the measurement of lux without the use of V-shaped book mounts, and within the first column of overhead lights only; the recorded value is 159 lux, which is well in excess of the recommended 50 lux for very sensitive materials. The authors note the comparison between the lux from overhead lights to the lux derived from four desk lamps (159 lux to 185 lux). The gain between these varying lighting conditions is 26 lux, which represents a significant increase above recommended values for extremely light-sensitive materials.

Special collections cannot be realistically illuminated to a 50 lux level, therefore the tangible value in collecting the raw data seen in table 3 is realized when it is used to calculate cumulative totals until a JNF occurs. The number of hours until a JNF occurs when collection material is accessed without a book mount was calculated using the formula and shown in table 4, allowing the steward to assess if the collection's materials are in danger of damaging light levels during routine use. By analyzing this data, alternative strategies can be planned for protecting the most valuable and sensitive materials.

With the lighting in the Ahmanson Murphy Reading Room accepted as the minimum required for task lighting, fading damage is estimated to occur on the most sensitive materials after 1,415 hours (140 days) or twenty-eight weeks of exposure. The most significant increase in lux is seen when one desk lamp is used in addition to the overhead lights, causing an increase of 30 lux, which translates to fading in 1,257 hours (126 days) or twenty-five weeks of exposure. Items that fall under the ISO categories of moderate and low sensitivities, as expected, show slower rates of fading, although they too are not immune to excessive light exposure, evidenced by the calculated hours to fading in table 4.

Table 3. Recorded Values of Consecutive Illumination

Light type	Overhead lights only	1 desk lamp	2 desk lamps	3 desk lamps (opposite side)	4 desk lamps (opposite side)
Values without book mounts					
Visible light (in lux)	159	179	181	184	185
Ultraviolet light (in microwatts per lumen)	45	43	43	43	43
Recorded values inside of V-shaped book mounts					
Visible light (in lux)	144	162	164	164	164
Ultraviolet light (in microwatts per lumen)	45	43	43	43	43

Table 4. Calculated Hours to Fading

ISO Blue Wool categories and hours to fading	Overhead lights only	1 desk lamp (reading directly under)	2 desk lamps	3 desk lamps (opposite side)	4 desk lamps (opposite side)
Calculated hours to fading based on light readings when books are flat					
High Sensitivity. Hours to Fading over 1.5 years	1,415	1,257	1,243	1,222	1,216
Moderate Sensitivity. Hours to Fading over 20 years	18,868	16,760	16,574	16,304	16,216
Low Sensitivity. Hours to Fading over 300 years	283,018	251,396	248,618	244,565	243,243
Calculated hours to fading, based on light readings taken within a V-shaped book mount					
High Sensitivity. Hours to Fading over 1.5 years	1,563	1,389	1,372	1,372	1,372
Moderate Sensitivity. Hours to Fading over 20 years	20,833	18,519	18,293	18,293	18,292
Low Sensitivity. Hours to Fading over 300 years	312,500	277,778	274,390	274,390	274,390

The authors recorded a second set of readings taken inside of foam V-shaped book mounts to assess if the use of these mounts affects the amount of light falling on collection material. The lux values from the second round of readings (table 3) showed a decrease in lux falling on collection material in comparison to the readings taken flat on the table. The recorded lux values were used in combination with the formula to derive the number of hours until JNF occurs. The calculated number of hours till JNF is also recorded in table 4.

With the overhead lights only, the lux reading was 144 when taken within V-shaped book mounts, in comparison to 159 lux without. While this is a difference of 15 lux, it translates to a difference of 147 hours (15 days) until JNF occurs. An unexpected benefit of the use of the V-shaped book mounts is the reduced light levels obtained in the readings, which can be seen by the consistent value of 164 lux recorded with two, three, and four sequential lights turned on. Comparing the hours to JNF between the readings taken with and without V-shaped book mounts revealed a dramatic reduction in lux levels. For clarity, only the ISO high-sensitivity materials are discussed. The lux levels taken while using V-shaped book mounts showed a JNF with only the overhead lights on occurring in 1,563 hours over 1.5 years (i.e., 142 days or 28 work weeks). This is a difference of fourteen days when compared to the hours until fading occurs without the V-shaped book mounts. The stable lux readings and the consequently static JNF shown with the use of two, three, or four sequential desk lamps turned on suggests that the book mounts not only protect the physical object, but also the information within.

The estimate of hours to fading within these studies does not account for previous fading and does not help users discriminate between materials of varying sensitivity. However, these exposure limits comply with the stringent AINSI/NISO standards of 420 days per year at 50 lux, that is, lighting in the special collection evaluated is three times as intense as the standard recommendation, so that the days to JNF are reduced to one-third. The ultraviolet radiation measured in the reading room is below the 75μW/lumen recommended by ANSI/NISO standards, but because it does not contribute to the visual access of these materials, the authors recommend filters that reduce the ultraviolet radiation to zero.

An unexpected finding of the measurements taken within V-shaped book mounts and with sequential lights turned can be seen with the reduction in light levels when compared to those readings measured when the books are flat. The finding of the reduced light levels with the use of V-shaped book mounts is unexpected because the authors had anticipated that an increase in light would result in an increase in recorded lux. The numerical difference in the results translates to a small extension of the lifespan of special collections materials. Knowing that one common preservation step taken at the Ahmanson Murphy Reading Room does not create risks in other areas is gratifying.

Conclusions and Recommendations

The primary research question for this study sought to assess the effects of variable illumination on special collections

during research use, the results of which have the potential to inform access, care, and library preservation policies. This study illustrates how recording lux hours may be used to determine rates of fading for collections of differing material sensitivities, information that may be used to devise concurrent access and preservation of materials. The assessment of the ambient and direct light within a reading room was conducted to evaluate the incident lux illuminating collection material during use, which is a strategy previously unreported in the literature. The measured values taken were used to calculate the total number of hours until a JNF would occur with the present lighting conditions. These calculations illustrate how various collection materials would fare using Michalski's guidelines for annual light exposure based on specific materials sensitivity.[46] These standards, designed for museum exhibition purposes, demonstrate that the risk of overexposure in a reading room setting is a concern only for the most fugitive and heavily used special collections materials. For these materials, information about lighting conditions and hours until fade may be used to establish an allowable number of use hours per year. However, use times typically required to correlate to a loss of information are not likely to be reached.

The approach taken in this study aimed to quantitatively evaluate the effects of variable illumination on special collections materials and found that the use of adjacent and opposing light sources contributes to the exposure of collection material. In this particular scenario, the variable lighting contributed to a minor increase in light exposure. The use of V-shaped book mounts provides support to collection material and deflects the angle of illumination, thereby diminishing the effects of deterioration by light damage. However, illumination will vary from institution to institution, which points to the need to know what the maximum lux potential can be per institution. Once that information is established, appropriate measures can be taken when very sensitive materials need to be accessed. In this study, the authors applied a museum conservation approach with the full understanding that ideal lighting for very sensitive materials may not always be possible. However, the ability to measure and calculate the lux in a given institution will ensure that access to information will remain possible, which is a concern for both librarians and conservators.

The authors' recommendations for special collections libraries include the implementation of data loggers to record light levels in spaces in which collection materials are accessed, and to collect incident readings simulating patron use to determine whether harmful light levels are present. The formula provided in this study enables collection stewards to evaluate lighting risks and take preservation action where needed. The authors suggest that a light log that records use hours be created and even be maintained specifically in the case of a patron requesting access to a significant subcollection, for example, illuminated and palm manuscripts, scrapbooks, artists' books, early newspapers, photographic media, and original art works within a special collection. These use hours plus the predetermined incident lux and ultraviolet radiation dose can be factored to calculate cumulative dose. The designation of lux hours per year may be used to consider access limitations to cumulative light energy for different collections' material sensitivities.

References

1. Meg Lowe Craft and M. Nicole Miller, "Controlling Daylight in Historic Structures: A Focus on Interior Methods," *APT Bulletin* 31, no. 1 (2000): 53–57.
2. Stefan Michalski, Ten Agents of Deterioration: Light, Ultraviolet and Infrared: How Much Light Do We Need to See? (Ottawa: Canadian Conservation Institute, 2010), www.cci-icc.gc.ca/crc/articles/mcpm/chap08-eng.aspx#Light (accessed Oct. 9, 2010).
3. Ibid.; Bertrand Lavédrine, *A Guide To the Preventive Conservation of Photograph Collections* (Los Angeles: Getty Conservation Institute, 2003): 152.
4. Lavédrine, *A Guide*, 152.
5. Ibid., 165.
6. Jane Henderson, *Managing the Library and Archive Environment* (London: National Preservation Office, 2007): 1.
7. American Institute for Conservation of Historic and Artistic Works, Caring For Your Treasures: Documents and Art on Paper, www.conservation-us.org/index.cfm?fuseaction=Page.ViewPage&PageID=628 (accessed Oct. 10, 2010); Stacy Etheredge, "Practicing Law Librarianship: Preserving a Special Collection," *AALL Spectrum* 11, no. 4 (Feb. 2007): 8–19, www.aallnet.org/products/pub_sp0702/pub_sp0702_ProDev.pdf (accessed Oct. 10, 2010).
8. Edward P. Adcock, ed., with Marie-Thérèse Varlamoff and Virginie Kremp, "Environmental Control" in *IFLA Principles for the Care and Handling of Library Material*, International Preservation Issues no. 1 (Paris: IFLA Core Programme on Preservation and Conservation; Washington D.C.: CLIR, 1998), http://archive.ifla.org/VI/4/news/pchlm.pdf (accessed Nov. 13, 2010).
9. Barclay Ogden, "Collection Preservation in Library Building Design" (Libris Design Project, 2004), http://librisdesign.org/docs/CollectionPreservation.doc (accessed Oct. 10, 2010).
10. Ibid, 18.
11. Edward T. Dean, "Daylighting Design in Libraries" (Libris Design Project, 2005): 9, table 1, http://librisdesign.org/docs/DaylightDesignLibs.pdf (accessed Sept. 27, 2010).
12. David R. Beech, "How to Look after Your Collection—A Basic Guide" (paper presented at the 87th Philatelic Congress of Great Britain, Derby, July 8, 2005), www.maidenheadphilatelic.co.uk/index_files/yourcollection.htm (accessed Oct. 10, 2010).
13. Etheredege, "Practicing Law Librarianship," 9.
14. Adcock, ed., *IFLA Principles*, 8.
15. National Information Standards Organization (NISO), *ANSI/NISO Z39.79-2001 Environmental Conditions for Exhibiting Library and Archival Materials* (Bethesda, Md.: NISO,

2001): appendix A, 17.
16. Ibid.; David Saunders, "Ultra-Violet Filters for Artificial Light Sources," *National Gallery Technical Bulletin* 13 (1989): 61–68.
17. Jonathan Ashley-Smith, Alan Derbyshire, and Boris Pretzel, "The Continuing Development of a Lighting Policy for Works of Art and Other Objects at the Victoria and Albert Museum," in *13th Triennial Meeting, ICOM Committee for Conservation, Rio de Janeiro*, ed. Roy Vontobel, 3–8 (London: James & James, 2002); Karen M. Colby, "A Suggested Exhibition Policy for Works of Art on Paper," *Journal of the International Institute for Conservation—Canadian Group* 17 (1992): 3–11; Alan Derbyshire and J. Ashley-Smith, "A Proposed Practical Lighting Policy for Works of Art on Paper at the V & A," in *12th Triennial Meeting, ICOM Committee for Conservation*, ed. David Grattan, 38–41 (London: James & James, 1999).
18. National Information Standards Organization, *ANSI/NISO Z39.79-2001*, 5–7; Library of Congress, Lighting of Library Materials (Aug. 10, 2010), www.loc.gov/preserv/care/light.html (accessed Sept. 26, 2010).
19. NISO, *ANSI/NISO Z39.79-2001*, 5–7; the directive (NARA 1571) is described in John Carlin, Archivist of the United States, "Archival Storage Standards," memorandum to Office Heads, Staff Directors, ISOO, NHPRC, OIG, National Archives and Records Administration, College Park, Md., February 15, 2002, p. 9.
20. Ogden, *Collection Preservation*, 18.
21. Mary Ellen Starmer, Sara Hyder McGough, and Aimée Leverette, "Rare Condition: Preservation Assessment for Rare Book Collections," *RBM: A Journal of Rare Books, Manuscripts, and Cultural Heritage* 6, no. 2 (2005): 91–107, http://rbm.acrl.org/content/6/2/91.full.pdf (accessed Sept. 27, 2010).
22. Lynn M. Thomas, "Special Collections and Manuscripts," in *Encyclopedia of Library and Information Sciences*, 3rd ed., ed. Marcia J. Bates and Mary Niles Maack, 4948–55 (London: Taylor & Francis, 2010).
23. Association of College & Research Libraries (ACRL), Guidelines on the Selection and Transfer of Materials from General Collections to Special Collections, 3rd ed. (approved by the ACRL Board of Directors, July 1, 2008): 4.0 Transfer Criteria, www.ala.org/ala/mgrps/divs/acrl/standards/selctransfer.cfm (accessed Nov. 13, 2010).
24. UCLA Library, Charles E. Young Research Library Department of Special Collections, Access to Special Collections Materials, Oct. 19, 2010, www.library.ucla.edu/specialcollections/researchlibrary/9619.cfm#access (accessed Nov. 13, 2010).
25. Ibid.
26. Lavédrine, *A Guide*.
27. Ibid.
28. Michalski, Ten Agents of Deterioration.
29. Beth Lindblom Patkus, Preservation Leaflets, 2.4, The Environment: Protection from Light Damage (Northeast Document Conservation Center, 2007), www.nedcc.org/resources/leaflets/2The_Environment/04ProtectionFromLight.php (accessed May 22, 2010).
30. Paul N. Banks, "Environment and Building Design," in *Preservation Issues and Planning*, ed. Paul N. Banks and Robert Pilette, 114–44 (Chicago: ALA, 2000).
31. Robert L. Feller, "The Deteriorating Effect of Light on Museum Objects," *Museum News Technical Supplement* 42, no. 3 (1964): 1–8.
32. Ann Garrison and William Lull, *Mechanisms of Environmental Damage to Collections* (Princeton Junction, N.J.: Garrison/Lull, 1996): 8.
33. David Saunders, "Protecting Works of Art From The Damaging Effects of Light," in *International Symposium on the Conservation and Restoration of Cultural Property, Tokyo National Research Institute of Cultural Properties* (Tokyo: Tokyo National Research Institute for Cultural Properties, 1990): 167–78.
34. Lavédrine, *A Guide*.
35. Michalski, Ten Agents of Deterioration.
36. Sarah Staniforth, "The Logging of Light Levels in National Trust Houses," *Preprints of the 9th Triennial Meeting, ICOM Committee for Conservation, Dresden, Germany, 26–31 August 1990* (Marina del Rey: Getty Conservation Institute, 1990): 602–6.
37. Jennifer E. Hain, "A Brief Look at Recent Developments in the Preservation and Conservation of Special Collections," *Library Trends* 52, no. 1 (Summer 2003): 112–17.
38. Patricia Morris, "Achieving a Preservation Environment with Data Logging Technology and Microclimates," *College & Undergraduate Libraries* 16, no. 1 (2009): 83–104.
39. Museums, Libraries, and Archives Council, *Benchmarks in Collection Care for Museums, Libraries and Archives: The Database Version* (2006): 97, www.collectionslink.org.uk/index.cfm?ct=assets.assetDisplay/title/Benchmarks%20in%20Collections%20Care%20for%20Museums,%20Libraries%20and%20Archives:%20the%20database%20version/assetId/7 (accessed Oct. 10, 2010).
40. Michalski, Ten Agents of Deterioration, table 4.
41. Robert Feller and Ruth Johnston-Feller, "The International Standards Organization's Blue-Wool Fading Standards (ISO R105)," in *Textiles and Museum Lighting* (paper presented at the Harper's Ferry Regional Textiles Group meeting, Washington, D.C., 1985): 41–57.
42. Ibid; Mauro Bacci et al., "Calibration and Use of Photosensitive Materials For Light Monitoring in Museums: Blue Wool Standard 1 as a Case Study," *Studies in Conservation* 49, no. 2 (2004): 85–98.
43. David Saunders and Jo Kirby, "Light-Induced Damage: Investigating the Reciprocity Principle," in *ICOM Committee for Conservation, 11th Triennial Meeting, Edinburgh, Scotland,* ed. Janet Bridgland, 87–90 (London: James & James, 1996).
44. Patkus, Preservation Leaflets, 2.4, The Environment: Protection from Light Damage.
45. Michalski, Ten Agents of Deterioration, table 3.
46. Ibid., table 5.

HathiTrust

A Research Library at Web Scale

Heather Christenson

Research libraries have a mission to build collections that will meet the research needs of their user communities over time, to curate these collections to ensure perpetual access, and to facilitate intellectual and physical access to these collections as effectively as possible. Recent mass digitization projects as well as financial pressures and limited space to store print collections have created a new environment and new challenges for large research libraries. This paper will describe one approach to these challenges: HathiTrust, a shared digital repository owned and operated by a partnership of more than forty major libraries.

> The activities of research libraries in the next five to 10 years will define the role of libraries in the digital age. The library community must now ensure that these collections not only retain their research value in a digital platform, but also realize their potential as users adjust their information needs and expectations.
>
> —HathiTrust FAQ, July 2010 (www.hathitrust.org/faq)

In an era of mass digitization of library collections, research libraries are confronting an array of new challenges to continuing their traditional role as stewards of library collections. How will libraries ensure perpetual preservation of these sometimes massive new digital library collections, a promise Google does not make? How will libraries provide wide access to their digital collections in an appropriate manner, unbeholden to commercial interests and in support of the activities of scholars? What new possibilities for services are opened up by digital formats, and how can libraries bring those new services to their user communities? How do these new large digital collections relate to print collections, and what opportunities are available for libraries to coordinate collection management between print and digital materials? This paper will consider these challenges and then describe how HathiTrust, a shared digital repository owned and operated by a partnership of more than forty major research libraries, offers answers to some of these questions and an opportunity for libraries to collectively explore this new territory.

Literature Review

Simultaneous with lively reporting and debate in prominent popular news sources and magazines regarding Google Books and the outcomes of mass digitization projects, researchers have explored the implications of mass digitization for libraries and the collaborative possibilities for addressing the challenges of digital preservation, access, support for scholarly research, and collection management in light of new, massive digital collections.[1] The specter of commercial hosting of research library content by Google juxtaposed with the responsibility of libraries to uphold their users' right to access information, as well as their

Heather Christenson (heather.christenson@ucop.edu) is Project Manager, Mass Digitization, California Digital Library, University of California, Oakland.

Submitted August 4, 2010; tentatively accepted September 3, 2010, pending modest revision; revision submitted November 2, 2010, and accepted for publication.

mission to preserve it, is a theme addressed by a number of researchers and library leaders. Hahn concluded that "it may be foolish to expect that commercial companies will share librarians' values and commitment to digitized material preservation" and that "research libraries alone will be held accountable for fulfilling that vital preservation mission."[2] In 2008, Brantley, then executive director of the Digital Library Federation, urged libraries to "trade for our own account" because libraries "stand for what no other organization in this world can: the fundamental right of access to information, and the compulsion to preserve it for future generations."[3] Leetaru made a case that the output of mass digitization is "access digitization" rather than "preservation digitization."[4] He acknowledged that placing responsibility for long-term storage with libraries "is a legitimate argument, especially in light of Microsoft's recent withdrawal from book digitization," but concluded that the academic community has so far failed to provide good access service for mass digitized books.[5] Dougherty also explored the question of what happens if Google goes away and pointed to HathiTrust as an example of libraries taking this question seriously.[6]

The utility of collaboration and scale for addressing problems of access and scholarly use of the mass digitized corpus is an idea that resonates with researchers. The Council on Library and Information Resources (CLIR), among others, has invested in moving research forward on the outcomes of mass digitization projects, and a number of CLIR-sponsored reports have been produced to this end. A 2007 report described ideas originating from a seminar on promoting digital scholarship and the "so-called 'million books' problem."[7] The report examined characteristics of the mass digitized corpus compared to local digitization methods (as they existed at the time), such as the greater heterogeneity of collections included, the variability of error rates that occur because of the optical character recognition (OCR) used in mass projects across texts and languages, and the lack of granular markup for logical pieces of text (e.g., chapters and sections, proper names). The report pointed to a potential model that combines "massive scale with the flexibility for particular domains to manage data and provide services that suit their needs."[8] In a 2008 paper, Rieger, addressing mass digitization projects, examined the "issues that influence the availability and usability, over time, of the digital books that these projects create," and recommended a balance of preservation and access requirements as well as collaboration amongst cultural institutions.[9] The concept of leveraging collaboration for cost savings in the development of repositories was mentioned by Furlough as he surveyed the repository landscape from a user-services perspective: "If content management and delivery services have a limited audience on a given campus, it may be better to partner with others."[10]

The implications of a library-owned aggregation of mass digitized materials for managing print and digital collections at the local institution level have emerged as a research theme in recent years. Sandler, in a thoughtful article, considered "a world where a single digital copy of an article or book can be delivered to multiple users, anytime, anywhere" and speculated that core resources could be "served up centrally," saving costs to individual libraries and enabling them to focus on needs specific to their institutions and user communities.[11] In the conclusion of a 2010 report, Henry pointed to a new collaborative cloud library model for collection development and management in which multiple libraries share the costs of maintaining both print books and their digital surrogates.[12] A recent project led by Malpas of OCLC Research and funded in part by a grant from the Andrew W. Mellon Foundation explored the proposition that outsourcing management of portions of monographic print collections because of replication in both shared digital and shared print storage may be cost effective for libraries.[13]

Context

The volume of books digitized from library collections has grown rapidly during the last decade. Output from small-scale, in-house library scanning operations was dwarfed when Google initiated its project to digitize books from libraries, first announced in December 2004. Google's stated goal for the Google Books Library Project is to "make it easier for people to find relevant books—specifically, books they wouldn't find any other way such as those that are out of print—while carefully respecting authors' and publishers' copyrights. Our ultimate goal is to work with publishers and libraries to create a comprehensive, searchable, virtual card catalog of all books in all languages that helps users discover new books and publishers discover new readers."[14] The Google Books Library Project was followed by the Open Content Alliance (OCA), a coalition of libraries, nonprofit organizations, and corporations formed in October 2005 with the goal of digitizing public domain works.[15] While a member of the OCA, and via its Live Search Books program, Microsoft funded the digitization of more than 750,000 books from libraries from December 2006 to May 2008 via its Live Search Books program.[16] By early 2008, the number of books digitized under the auspices of these programs began to approach many millions across the participating libraries.

With libraries facing enormous economic pressures and with Google's projects to digitize libraries' collections continuing to increase the amount of digitized content, a number of research libraries joined together to address these issues. In October 2008, HathiTrust was launched as a collaborative effort by the Committee on Institutional Cooperation (CIC)[17]—then a consortium of thirteen universities, two of which (Michigan and Wisconsin) were already

Google Library partners—and the University of California Libraries to create a shared repository of digital collections.[18] The University of Virginia became a participant in January 2009, with many other libraries joining since then. These partners have joined with the common understanding that the massive scale of library digitization enterprises, along with the high costs of digital preservation, demand a web-scale collaborative solution for ensuring long-term access to the digital output and a new vision for a collective collection. Because of the size of the HathiTrust repository and the depth of the collaboration involved, the participating libraries are uniquely positioned to leverage technical infrastructure and collective expertise for digital preservation, services, and collection management on an unprecedented scale. The presence of a critical mass of research institutions in the HathiTrust partnership enables an aggregation of digital resources not seen before, hosted by libraries for the long term in a continuation of their traditional role as stewards of the scholarly record and supporters of research and other scholarly pursuits.

What is HathiTrust?

At the heart of HathiTrust is a shared secure digital repository owned and operated by a partnership of major research libraries. The repository is best known as a means of preserving digital materials created via large-scale digitization projects. By pooling their collective resources and expertise, the partners have created a robust and scalable infrastructure to efficiently store, manage, and preserve their collections of digital books and journals in common. HathiTrust, however, has positioned itself as more than a simple aggregation of digitized library material. Its stated mission is much broader: "to contribute to the public good by collecting, organizing, preserving, communicating, and sharing the record of human knowledge."[19]

To that end, the HathiTrust repository now likely contains the largest collection of digital volumes outside of Google Books. Because most of the U.S.-based Google library partners are members, the collections of the current HathiTrust members can be estimated to constitute a majority of all of the content contributed by U.S. libraries to Google Books. The partnership is open to institutions internationally. The first partner from outside the United States is the Universidad Complutense de Madrid, also a Google Books library partner. As HathiTrust adds members, the repository also will encompass a growing number of the volumes digitized from U.S. libraries by Microsoft under the auspices of the now defunct Microsoft Live Search Books service. In addition, the repository now contains tens of thousands of volumes digitized by the Internet Archive and additional volumes digitized by the partners themselves. Growth has been rapid, and the repository (as of this writing) holds more than 8 million volumes, including 2 million public domain volumes.[20] As the combined output of mass digitization accumulates, the sheer number of digital volumes aggregated in the repository will foster the partner libraries' collective ability to leverage a digital version of library collections assembled and curated by generations of librarians across the nation's research libraries. In addition to including the digital volumes derived from print, plans are underway to include other types of digital publications within the repository. For example, HathiTrust is in discussions with university presses about putting new books and book backlists online via open access and plans to extend that model. Currently, hundreds of current university-press titles are available online with open access permissions. The partners intend that the repository eventually will encompass materials beyond books and journals. Because new content types will demand new access, management, and preservation requirements, much remains to be resolved.

Goals and Values

The name HathiTrust was chosen to express the fundamental values of the organization. Hathi (pronounced hah-tee) is the Hindi word for elephant, an animal noted for its memory, wisdom, and strength. While HathiTrust's intent is to build a reliable and increasingly comprehensive digital archive of library materials converted from print that is co-owned and managed by a number of research institutions, the enterprise has a number of other important goals:

- To dramatically improve access to these materials in ways that, first and foremost, meet the needs of the co-owning institutions.
- To help preserve these important human records by creating reliable and accessible electronic representations.
- To stimulate efforts to coordinate shared collection management strategies among libraries, thus reducing long-term capital and operating costs of libraries associated with the storage and care of print collections.
- To create and sustain this "public good" in a way that mitigates the problem of free riders.
- To create a technical framework that is simultaneously responsive to members through the centralized creation of functionality and sufficiently open to the creation of tools and services not created by the central organization.[21]

HathiTrust differs from Google and from organizations such as the Internet Archive in a number of ways. Structurally, HathiTrust is not a corporation or even a nonprofit organization nor is it a "trust" in the legal sense of the word. The

partnership is a collaborative enterprise of research libraries that depends on funding and in-kind contributions from members. As "an enterprise principally driven by a scholarly mission," HathiTrust is committed to the principle that "creating a digital research library for the research community is the responsibility of research libraries."[22] In accordance with its research mission, HathiTrust embraces values long held by libraries, such as preservation, quality, privacy, and public access, and formally commits to long-term digital preservation. Although Google and the Internet Archive both maintain and provide access to large amounts of data as a matter of course, neither organization is formally committed to digital preservation of digitized books over time. HathiTrust also differs from national or regional projects such as the Joint Information Systems Committee (JISC) (www.jisc.ac.uk) in the United Kingdom or the Europeana (www.europeana.eu/portal) initiative of the European Union in that currently no government-supported mandate or national cultural institution supports its existence.

In keeping with a public access mission, HathiTrust has put mechanisms in place to support greater access to the works in the repository. Although HathiTrust must follow copyright law and restrict access to volumes that are not in the public domain, the organization's philosophy is to open up materials to the greatest extent legally permissible. Most of the digital volumes within the repository are the result of Google's digitization, but HathiTrust may assign a viewability status to the library copy that is different from that of the copy in Google Books. In general, HathiTrust takes a less conservative stance regarding providing full-view access to government documents. In addition, a growing number of HathiTrust institutions provide mechanisms for rights holders to release their works into full view within the HathiTrust.

HathiTrust partner libraries also are actively working to move orphan works (works that can be assumed to be in-copyright but whose copyright owner cannot be located) into the public domain as another route to greater access. As of this writing, 6 million volumes within the repository are considered in-copyright or potentially in-copyright orphan works and are not viewable, but Lavoi and Dempsey note that many of those fall into the orphan works category and actually may be in the public domain.[23] By collaborating within HathiTrust, research libraries plan to begin tackling this problem collectively. With support from an Institute of Museum and Library Services (IMLS) National Leadership grant, the University of Michigan has developed a Copyright Review Management System (CRMS), the expansion of which is currently being piloted by several HathiTrust partner libraries under the aegis of the IMLS grant.[24] This system will be a means to scale and propagate book-by-book copyright determination by human beings, a process that can be arduous and complex. The intent is to expand use of the CRMS for copyright review activities across research libraries. In the meantime, Michigan is making progress, having used the CRMS to analyze more than 123,000 books (as of this writing) and moved approximately 54 percent of them into the public domain. HathiTrust makes these rights determinations available as part of a set of downloadable metadata called the "Hathifiles."[25]

Any discussion of copyright and digitized books invariably leads to the Google Books Settlement Agreement. The October 2008 agreement between the Authors Guild, the Association of American Publishers, and Google settled *Authors Guild et al. v. Google*, a class-action lawsuit alleging that Google's digitization and indexing of in-copyright works constitutes copyright infringement.[26] In November 2009, an amended version was filed in response to a Department of Justice brief suggesting that the original version violated antitrust laws. The amended settlement is complex and has engendered discussion much broader than the scope of this paper, and, of this writing, the presiding judge has yet to rule. The aims of HathiTrust predate and are independent of the settlement and the amended settlement. However, HathiTrust could be affected by some of the provisions of the amended settlement, including those that would allow Google to sell institutional subscriptions to libraries for full view of books within Google, provide for libraries to host a research corpus of books, and prescribe the establishment of a "book rights registry." In a December 2008 interview, John Wilkin, HathiTrust Executive Director, addressed some of the more positive potential effects:

> Much of what HathiTrust proposes to do—preserve content, support access by print-disabled users, generate print replacement copies from the digital files when original print copies are damaged or lost, and serve as a body of content for large-scale computational needs—is explicitly sanctioned in the settlement agreement, thus protecting this fundamental library-based effort from legal threats.[27]

HathiTrust service development will need to take the settlement outcomes into account, respecting mandated constraints where they exist. If the amended settlement is approved, HathiTrust may leverage services such as the institutional subscription within the access services it offers where appropriate and considered valuable to the partners.

Collaboration

Owning and managing the repository is of inherent benefit to the participating libraries, and such an enterprise demands a thoughtfully structured collaborative infrastructure that accounts for the interests of all partners. In addition to cost savings for digital preservation and services resulting from economy of scale, a key benefit of collaboration is the ability

to tap into expertise across the libraries. HathiTrust has a shared governance structure, with an executive committee that is the decision-making body, along with a strategic advisory board composed of university librarians and associate university librarians from the partner institutions. The strategic advisory board sets functional objectives, convenes task forces to address specific issues, and recommends policies, drawing on the array of experience and expertise of the members.

Within the past year, HathiTrust launched working groups on a wide range of topics, including communications, collection development and management, quality, ingest and error rates, collaborative development, resource discovery, faculty research, and storage expansion needs. In addition to these formal groups, HathiTrust has brought together technical talent from the participating institutions to develop and improve its operational processes and aims for more collaborative development in the future. Both the organization and individual participants are gaining experience in long-term collaboration on core infrastructure and services. In a stringent economic climate in which libraries are increasingly seeking to collaborate, the growing pool of expertise gained by participants becomes a valuable asset.

Preservation

Secure and long-term digital preservation of volumes in the repository is fundamental to the goals of the enterprise. The HathiTrust repository is sometimes compared to the Portico (www.portico.org) and CLOCKSS (Controlled Lots of Copies Keep Stuff Safe) (www.clocks.org) digital preservation services, but it differs from them in terms of the provenance of included content, archival philosophy, and underlying business and organizational structure. Both Portico and CLOCKSS focus on journal and e-book content originating from publishers, while HathiTrust has begun with content from the libraries' mass digitization projects. Both Portico and CLOCKSS are dark archives that make content available only when a trigger event (such as a publisher ceasing operation) occurs. Although a large amount of content within the HathiTrust repository is not viewable to end users by copyright law, all other content is available and the repository is technically a light archive. Portico and CLOCKSS are services of nonprofit ventures and partner with publishers, while HathiTrust is an organization composed solely of libraries.

HathiTrust is committed to preserving the intellectual content and, if reasonably possible, the exact appearance and layout of materials digitized for deposit and is committed to allowing the partners to make open and meaningful decisions about formats and quality. For example, upon joining, a partner institution may determine which image file format they want to use for their deposited content, and the decision process of each partner may be documented and shared to inform the others. Individual partner institutions may have varying positions on whether the digital copies of print books created via mass process are preservation-worthy copies, but HathiTrust is seeking to conduct research in this area and develop quality metrics. With funding from the Andrew W. Mellon Foundation, Paul Conway of the University of Michigan School of Information, in conjunction with HathiTrust, is investigating means of measuring quality and usefulness of digital objects and the feasibility of establishing a mechanism for branding the trustworthiness of deposited volumes for particular uses, such as reading, printing volumes on demand, and performing computational research.[28] The goal of this certification process is to "give assurance that content within a repository is worthy of preservation, and increase the value of that content in broader discussions about storage and management solutions for both digital and print collections."[29]

The HathiTrust repository conforms to accepted standards and models for digital preservation, including the International Standards Organization's Open Archival Information System (ISO OAIS) reference model, the Metadata Encoding and Transmission Standard (METS), and the Preservation Metadata Implementation Strategies (PREMIS) Data Dictionary.[30] Digital objects are stored in formats that are documented, open, and standards-based with the intent of providing an effective means to migrate objects to successive preservation formats over time, as necessary. The repository utilizes robust technology and has the geographic redundancy of two mirror sites at the University of Michigan and Indiana University. In addition, each site has several layers of redundancy; a tape backup constitutes yet another copy. A cross-institutional working group reviewed the storage configuration, conducted a cost–benefit analysis regarding the need for more redundancy, and reported a "high level of confidence in the existing two-instance architecture."[31] The Center for Research Libraries is now reviewing HathiTrust for Trustworthy Repositories Audit and Certification (TRAC) compliance.[32] The TRAC review is an independent evaluation that gauges a repository's capability to reliably store, migrate, and provide access to collections, and it is sought after by preservation repositories as a community metric of confidence.

By virtue of its scale and its acceptance of varied content from many different sources, the repository is well suited for encountering and overcoming common challenges, specifically in areas of repository standards, best practices, and methods for certifying the quality of the deposited content. During the past year, the creation of new content-ingest streams has tested the original repository structure built by the University of Michigan, reinforcing some principles for homogeneity of file formats and metadata and

also identifying where the partners can make choices and where flexibility is required. A team of members from the University of Michigan and the California Digital Library (on behalf of the University of California (UC)) collectively tackled the creation of two new content-ingest streams: UC's Google-digitized volumes and UC's Internet Archive-digitized volumes. The group faced the technical challenges associated with allowing heterogeneity and ensuring the ability of the repository and its services to function. As each content stream was created, the team gave rigorous attention to choices about such elements as identifiers, image formats, individual files selected from digitization vendors' content packages, and specific tags and associated variables in the recording of the progression of transformative events upon ingest within preservation metadata. Effective management of objects in the repository must encompass digital preservation standards and uses within access services. Accommodating these dual purposes can present a technical challenge. For example, a choice may be made that PDF is not appropriate for preservation, although that format may be useful for end-user access. In that case, the object may be stored in a more preservation-appropriate format and, for access purposes, the PDF format may be derived from the preservation file format, requiring an extra process on the access end, a compromise that serves both purposes.

Discovery and Access

In addition to digital preservation and in concert with it, HathiTrust embraces access services as essential to its mission. The HathiTrust repository offers a number of end-user services, such as basic and advanced bibliographic search, full-text search based on extracted text, and a collection-builder tool (explained below). The bibliographic search uses an aggregation of records contributed by partner libraries and thus is based on rich descriptive metadata that is the output of decades of library cataloging. The bibliographic search is comprehensive across the full spectrum of the digital collection from in-copyright to public domain. Researchers have documented Google's metadata errors, primarily resulting from automated processes.[33] Since HathiTrust metadata originates from partner libraries, the libraries have a more direct opportunity to resolve errors, collectively explore how the original cataloging of print volumes can be enhanced and extended to digital volumes, and experiment with optimally integrating bibliographic metadata with full text for search purposes. The full-text search (also known as large-scale search) was built by developers at the University of Michigan, and further development is guided by the HathiTrust Discovery Interface working group. As of this writing, the repository is providing full-text search across more than 2.8 billion pages contained within 8 million volumes. A distinctive feature of this service is that libraries own both the search mechanism and the content on which it acts. This ownership is significant for several reasons. For end users, the selectiveness and ranking of search results are not influenced by commercial interests, and the material covered by the search is a known corpus of materials selected, cataloged, and curated by librarians with the interests of academic users in mind. For partner libraries, owning the full-text search and the content provides an opportunity to engineer end-user services that are configurable for scholarly uses as well as free from advertising, commercial bias, and censorship.

Once discovered, digital volumes within the repository are accessible by various means depending on copyright status. Google-digitized public domain volumes are available in a full PDF download to authenticated users from partner institutions; public domain volumes digitized via Internet Archive and locally by partners are available in full PDF to all. All public domain volumes can be viewed on the web in a page-turner application. Volumes that are in copyright are discoverable via large-scale search, and users may view a list of pages on which their search term appears (snippets are not yet available). Most books are treated as in-copyright, but may be moved to an open status upon human-reviewed copyright determination (e.g., through the CRMS or on request from the rights holder). HathiTrust also offers services to print-disabled users who are located at the University of Michigan and plans to extend the service to other partners.[34] Printed versions of public domain books from some partners are now offered via a link within the HathiTrust Interface to print-on-demand service.

The Collection Builder functionality allows librarians and individual end users to create and share specific themed collections regardless of whether the end user is affiliated with a partner institution. The Collection Builder has great potential for integration within local services, such as online courses and themed collection portals built by local institutions. Once a collection is created, the full text of those volumes can be searched as a set. One can envision other future scholarly tools that can capitalize on a scoped, curated group of volumes by being able to manipulate and analyze them in various ways.

In keeping with its mission to enable local institutions to develop tools and services, the HathiTrust offers freely available data, open to any institution, that can be captured and incorporated in a local service. The data also is machine-accessible so that local services can be built using it. For example, the University of California uses data to provide direct links to the full text of HathiTrust public domain volumes via UC-eLinks (www.cdlib.org/services/d2d/ucelinks), its local link resolution service. A growing number of partner libraries provide links to HathiTrust resources within their online public access catalogs.

Supporting Research

Also emerging is support for scholarly computational research. During the past year, a working group convened to develop specifications for a research center for scholarly use. This action was taken in anticipation of the pending Google settlement, which includes terms that sanction the use of in-copyright works owned by HathiTrust institutions in "non-consumptive" computational research. Non-consumptive research is understood to describe "analysis of a form that does not require (and does not permit) reading access to in-copyright materials."[35] The terms of the settlement also provide for the establishment of up to two research centers that would enable this research across the entire body of Google-scanned content. HathiTrust is proposing a center that will support research capabilities across the HathiTrust corpus, which it defines as "the complete set of works in HathiTrust, including Public Domain, Google Public Domain, Open Access, and In-copyright Data."[36] The report states,

> The founding institutions of HathiTrust undertook the effort of building a repository of published content with the expectation that this content in addition to serving the needs of traditional reading and research would serve as an extraordinary foundation for many forms of computing-intensive research, particularly in the areas of language and literature.[37]

The working group characterized research types that a HathiTrust research center would need to support, including aggregation and distillation of subsets of data, development of tools, mechanisms for collaboration, and ability to preprocess and add data. Using this collectively defined framework, HathiTrust has begun to investigate the Software Environment for the Advancement of Scholarly Research (SEASR) (http://seasr.org) as a means to provide computational access to materials stored in the repository. SEASR, funded by the Andrew W. Mellon Foundation, is a research and development environment devoted to supporting digital humanities initiatives and fostering collaboration in a virtual environment.

Collection Development and Management

Underlying these services is the HathiTrust collection. The numbers do not tell the whole story of the depth and breadth of the collection; however, numbers give a frame of reference and starting point. At the time of this writing, within the more than 8 million total volumes (2 million volumes in the public domain), 4.5 million book titles and nearly 200,000 serial titles are represented.[38] The current HathiTrust collection spans several centuries and hundreds of languages. The top ten languages (English, German, French, Russian, Chinese, Spanish, Japanese, Italian, Arabic, and Polish) account for approximately 86 percent of the content, and the next forty languages account for another 13 percent.[39] U.C. Berkeley University Librarian Thomas Leonard has commented that we can view the HathiTrust collection in the same way astronomers look far out into the universe; like the images of stars that are light years away and thus ancient, the further back we go into the collection, the more we see a snapshot of what research libraries were collecting at the time.[40]

An analysis performed by Malpas on the subject distribution of titles, based on subject headings within bibliographic metadata, revealed "Language, Linguistics, and Literature" and "History and Auxiliary Sciences" to be the most populous subjects, followed by "Business and Economics," "Philosophy and Religion," and "Art and Architecture."[41] The HathiTrust website provides visualizations of the collection categorized by Library of Congress classification, language, and publication date.[42] Analysis of bibliographic metadata is only beginning to explore the types of collection analysis that might be possible via the full text search and specialized tools. Having bibliographic metadata, digital content, and management metadata in a common repository under library ownership likely will foster the development of analysis tools to answer questions that cross the boundaries of the data and depend on the synergy of the aggregation. For example, what has been collectively digitized, and what format is it in? What is the array of conditions that create a true duplicate, how much duplication is present, and what de-duplication strategies make sense?

HathiTrust formed a collections committee that will explore what additional tools and services may be needed to characterize the collection as it evolves. These may include analytical tools that examine subject, language, date, format, or other characteristics; extensions of the Collection Builder tool; and mechanisms that would be useful to describe the corpus to a potential user.

In concert with those activities, the HathiTrust corpus can be used as a basis for the development of comprehensive or distinctive digital collections in particular areas that build on participant strengths. The collections committee will tackle those opportunities as well. For example, the partners could develop a shared approach to government documents that capitalizes on the CIC's focused U.S. government documents digitization initiative.[43] Gap analysis and collection building will likely lead the partners to explore opportunities for digitization and collaboration with other initiatives.

Print Curation

Leveraging the HathiTrust corpus to manage print

collections both within and beyond the partner libraries is an active area of exploration. Driven by economics and space constraints, momentum is building toward putting ideas about collective print curation into practice. The mechanics of how aspects of this might work are beginning to emerge through recent research.

Malpas' 2010 Cloud Library research project explored the proposition that outsourcing management of portions of monographic print collections, based on replication in both shared digital and shared print storage, may be cost effective for libraries.[44] The study revealed marked overlap between the HathiTrust monographic collection and the holdings of major shared print repositories across the country, and thus a large potential library clientele for outsourced service. The study also found that until in-copyright works can be distributed digitally, the tipping point for cost-effectiveness would likely not be reached for most libraries. In addition, the libraries of the CIC universities have undertaken federal government documents digitization with an eye toward examining the relationship between print and digital copies to better position themselves for coordinated decisions about print retention.[45] Although small steps, these two examples, along with a trend toward shared print storage initiatives evidenced in discussions at the April 2010 meeting of the Association of Research Libraries, can be seen as early indicators of what is to come and of the economic incentives and collaborative structures that may be needed.[46]

HathiTrust has recently developed a cost model for participation to include libraries that may wish to leverage the digital collection for print collection management and other purposes.[47] The initial participation model has been that institutions pay infrastructure costs for the digital content they contribute. The second, newer participation model is aimed at institutions that do not necessarily have large collections (or any) of digital content to contribute but want to participate in the curation and management of the repository in return for specialized services. By paying a membership fee, these partners will contribute to sustaining a common resource, share in uses of relevant materials, and have a voice in future directions of HathiTrust. The second model also addresses the problem of "free riders," avoiding a situation where some partners would have access to an amount of content out of proportion to the amount of their monetary contribution. The newer participation model is based on partners' print holdings, and costs are calculated on a number of precise elements about cost and the "sharedness of the content," including costs to maintain public domain content.

Dempsey has used HathiTrust as an example of how "web scale" activity is "managed at the network level," and its "audience is potentially all web users."[48] Although all members pay, the network level of the HathiTrust infrastructure enables the libraries to pool their resources and reach more users more effectively at lower costs and to effectively "transfer resource[s] away from 'infrastructure' and towards user engagement."[49] The general public also benefits from this arrangement through access to public domain resources and discovery services.

Next Steps

Technical challenges are perhaps the easiest for pioneering organizations to overcome. Much more difficult are the challenges of achieving collaboration and political harmony, agreeing on policy, and implementing and building a new organizational culture within a group of geographically dispersed institutions with independent governance structures. In light of these challenges, HathiTrust has a plan for the next steps of its evolution. Following a formal review of the repository by partners, HathiTrust will convene its first Constitutional Convention in 2011. At this convention, the partners will have an opportunity to enhance, revise, or reenvision governance, partnership, and cost models.

Looking toward the future, the membership will need to continue to think boldly. Could the HathiTrust's mantle of stewardship and the values it embraces enable it to evolve into a broader role as a de facto national research library? Might even commercial agents (such as Google) come to view HathiTrust as a solution to the problem of long-term digital preservation? Even if these entities do realize that HathiTrust can fulfill the need for digital preservation, how do the public good aspects of HathiTrust's mission intersect with the commercial interests of for-profit enterprises? What new partnerships can be formed to advance the scholarly agenda of the HathiTrust partners? When research libraries collectively hold digital copies of significant portions of their collections, how comfortable will they be with collectively pushing the boundaries of legal use of the digital copies, and how effectively can they advocate for copyright reform?

Conclusion

In naming the founding of HathiTrust as one of *Library Journal*'s top academic library stories of 2008, Albanese described it as "the library community's most ambitious digital collaboration ever."[50] Two years later, the HathiTrust partners are making progress on issues such as cost-effective digital preservation of very large collections of digital volumes, access mechanisms for such a collection, including openly available metadata, and support for computational research. HathiTrust represents a growing digital aggregation of research library content at a scale with the potential to support collection management decisions as research libraries face financial pressures and weigh the relative value

of print and digital volumes. The widespread collaboration, aggregated expertise, and pooled digital collections of HathiTrust seem to be resulting in beneficial progress for both the library community and end users.

References and Notes

1. Jeffrey Toobin, "Google's Moon Shot: The Quest for the Universal Library," *New Yorker*, Feb. 5, 2007, www.newyorker.com/reporting/2007/02/05/070205fa_fact_toobin (accessed Oct. 8, 2010); Sergey Brin, "A Library to Last Forever," *New York Times*, Oct. 9, 2009, www.nytimes.com/2009/10/09/opinion/09brin.html (accessed Oct. 31, 2010).
2. Trudi Bellardo Hahn, "Mass Digitization: Implications for Preserving the Scholarly Record," *Library Resources & Technical Services* 52, no. 1 (Jan. 2008): 18, 24.
3. Peter Brantley, "Book Search Will Not Work Like Web Search," online posting, Jan. 1, 2008, Peter Brantley's Thoughts and Speculations, http://blogs.lib.berkeley.edu/shimenawa.php/2008/01/02/trade_for_our_own_account (accessed Oct. 1, 2010).
4. Kalev Leetaru, "Mass Digitization: The Deeper Story of Google Books and the Open Content Alliance," *First Monday* 13, no. 10 (Oct. 6, 2008), www.uic.edu/htbin/cgiwrap/bin/ojs/index.php/fm/article/viewArticle/2101/2037 (accessed Oct. 31, 2010).
5. Ibid.
6. William C. Dougherty, "The Google Books Project: Will It Make Libraries Obsolete?" *Journal of Academic Librarianship* 36, no. 1 (Jan. 2010): 86–89.
7. Council on Library and Information Resources, *Many More than a Million: Building the Digital Environment for the Age of Abundance. Report of a One-Day Seminar on Promoting Digital Scholarship Sponsored by the Council on Library and Information Resources (Nov. 28, 2008): Final Report* (Mar. 1, 2008), www.clir.org/activities/digitalscholar/nov28final.pdf (accessed Oct. 1, 2010).
8. Ibid.
9. Oya Y. Rieger, *Preservation in the Age of Large-Scale Digitization: A White Paper*, CLIR Publication 141 (Washington, D.C.: Council on Library and Information Resources, 2008): vi, www.clir.org/pubs/reports/pub141/pub141.pdf (accessed Nov. 11, 2010).
10. Mike Furlough, "What We Talk About When We Talk About Repositories," *Reference & Users Services Quarterly* 49, no. 1 (2009): 22.
11. Mark Sandler, "Collection Development in the ~~Age~~ Day of Google," *Library Resources & Technical Services* 50, no. 4 (Nov. 30, 2005): 241.
12. Charles Henry, "Epilogue," *The Idea of Order: Transforming Research Collections for 21st Century Scholarship*, CLIR Publication 147 (Washington D.C.: Council on Library and Information Resources, 2010): 121–23, www.clir.org/pubs/reports/pub147/pub147.pdf (accessed Sept. 2, 2010).
13. Constance Malpas, *Cloud-Sourcing Research Collections: Managing Print in the Mass-Digitized Library Environment* (Dublin, Ohio: OCLC Research, 2011), www.oclc.org/research/publication/library/2011/2011-01.pdf (accessed Jan. 19, 2011).
14. Google Books, Google Books Library Project—An Enhanced Card Catalog of the World's Books, books.google.com/googlebooks/library.html (accessed Sept. 3, 2010).
15. Microsoft, "Microsoft Live Search Fact Sheet," www.microsoft.com/presspass/newsroom/factsheet/LiveSearchFS.mspx (accessed Oct. 8, 2010).
16. Open Content Alliance, Global Consortium Forms Open Content Alliance to Bring Additional Content Online and Make It Searchable, web.archive.org/web/20051007010920/http://www.opencontentalliance.org/OCARelease.pdf (accessed Oct. 28, 2010).
17. Committee on Institutional Collaboration, About CIC, www.cic.net/Home/AboutCIC.aspx (accessed Oct. 31, 2010).
18. HathiTrust, Major Library Partners Launch HathiTrust Shared Digital Repository, www.hathitrust.org/press_10-13-2008 (accessed June 30, 2010).
19. HathiTrust, Mission and Goals, www.hathitrust.org/mission_goals (accessed Aug. 1, 2010).
20. HathiTrust, Currently Digitized, www.hathitrust.org/ (accessed Feb. 10, 2011).
21. HathiTrust, Mission and Goals, www.hathitrust.org/mission_goals (accessed Aug. 1, 2010).
22. Roger C. Schonfeld, "Conclusion," in *The Idea of Order: Transforming Research Collections for 21st Century Scholarship*, CLIR Publication no. 147 (Washington D.C.: Council on Library and Information Resources, 2010): 116–20, www.clir.org/pubs/reports/pub147/pub147.pdf (accessed Sept. 2, 2010); HathiTrust, "FAQ", www.hathitrust.org/faq (accessed July 21, 2010).
23. Brian Lavoie and Lorcan Dempsey, "Beyond 1923: Characteristics of Potentially In-Copyright Print Books in Library Collections," *D-Lib Magazine* 15, no. 2/1 (Nov./Dec. 2009), www.dlib.org/dlib/november09/lavoie/11lavoie.html (accessed Oct. 31,2010).
24. University of Michigan Library, Copyright Review Management System—IMLS National Leadership Grant, Copyright Review Management System, www.lib.umich.edu/imls-national-leadership-grant-crms (accessed Aug. 30, 2010).
25. HathiTrust, Hathifiles Metadata, www.hathitrust.org/hathifiles_metadata (accessed Aug. 3, 2010).
26. Jonathan Band, *A Guide for the Perplexed Part III: The Amended Settlement Agreement* (Washington, D.C.: ALA, Association of College and Research Libraries, and Association of Research Libraries, 2009), www.arl.org/bm~doc/guide_for_the_perplexed_part3_final.pdf (accessed Oct. 29, 2010).
27. John Wilkin, "BackTalk: HathiTrust and the Google Deal," *Library Journal* (Dec. 23, 2008), www.libraryjournal.com/article/CA6624782.html (accessed Oct. 31, 2010).
28. University of Michigan School of Information, "Mellon Grant Aids Research Criteria for Digital Libraries," online posting, Sept. 28, 2009, SI Informant, blog.si.umich.edu/2009/09/28/mellon-grant-aids-researching-criteria-for-digital-libraries (accessed Aug. 30, 2010).
29. HathiTrust, Update on October 2009 Activities, www

.hathitrust.org/updates_october2009 (accessed July 30, 2010).
30. ISO Archiving Standards—Reference Model Papers, http://nssdc.gsfc.nasa.gov/nost/isoas/ref_model.html (accessed Oct. 31, 2010); Library of Congress, Metadata Encoding and Transmission Standard, www.loc.gov/standards/mets (accessed Oct. 28, 2010); Library of Congress, PREMIS Data Dictionary for Preservation Metadata Version 2.0, www.loc.gov/standards/premis, (accessed Oct. 28, 2010).
31. HathiTrust, Recommendations Regarding a Third HathiTrust Instance, www.hathitrust.org/documents/hathitrust-3rd-instance-recommendations.pdf (accessed July 30, 2010).
32. OCLC and Center for Research Libraries, *Trustworthy Repositories Audit & Certification: Criteria and Checklist Version 1* (Chicago: CRL; Dublin, Ohio: OCLC, 2007), www.crl.edu/sites/default/files/attachments/pages/trac_0.pdf (accessed Aug. 1, 2010).
33. See, for example, Geoffrey Nunberg, "Google Books: A Metadata Train Wreck," online posting, Aug. 29, 2009, Language Log, languagelog.ldc.upenn.edu/nll/?p=1701 (accessed July 29, 2010).
34. Paul Courant, "Digitization and Accessibility," online posting, Nov. 2, 2009, Au Courant, paulcourant.net/2009/11/02/digitization-and-accessibility (accessed Aug. 1, 2010).
35. HathiTrust, Call for Proposal to Develop a HathiTrust Research Center, www.hathitrust.org/documents/hathitrust-research-center-rfp.pdf (accessed July 30, 2010).
36. Ibid.
37. Ibid.
38. HathiTrust, Statistics Information, www.hathitrust.org/statistics_info (accessed Feb. 10, 2011).
39. HathiTrust, HathiTrust Languages, www.hathitrust.org/visualizations_languages (accessed Aug. 31, 2010).
40. Thomas Leonard, "From Print to Digital: Visions of 21st Century Collections" (presentation, Pacific Rim Digital Alliance 2010 Meeting, Shanghai, China, October 21–22, 2010).
41. Constance Malpas, OCLC Research, "Subject Distribution of Titles in the Hathi Repository June 2010" (slideshow presentation at the American Library Association Annual Conference, Washington, D.C., June 2010), www.slideshare.net/oclcr/june-2010-subject-snapshots (accessed July 30, 2010).
42. HathiTrust, Visualizations, www.hathitrust.org/visualizations_callnumbers (accessed Aug. 1, 2010).
43. Committee on Institutional Collaboration, "CIC–Google Government Documents Project," www.cic.net/Home/Projects/Library/BookSearch/Govdocs.aspx (accessed Aug. 1, 2010).
44. Malpas, *Cloud-Sourcing Research Collections.*
45. Committee on Institutional Collaboration, "CIC-Google Government Documents Project."
46. Lizanne Payne, "Models for Shared Print Archives: WEST and CRL," (slideshow presentation at the 156th Association of Research Libraries Membership Meeting, Seattle, Washington, Apr. 28–30, 2010), www.arl.org/bm~doc/mm10sp-payne.pdf (accessed Nov. 11, 2010).
47. John Wilkin, memo, Feb. 12, 2010, www.hathitrust.org/documents/hathitrust-cost-rationale-2013.pdf (accessed Oct. 31, 2010).
48. Lorcan Dempsey, "Sourcing and Scaling," online posting, Feb. 21, 2010, Lorcan Dempsey's Weblog, orweblog.oclc.org/archives/002058.html (accessed July 29, 2010).
49. Ibid.
50. Andrew Albanese, "The Library Journal Academic Newswire Year in Review, the Top Academic Library Stories of 2008," *Library Journal* (Jan. 7, 2009), www.libraryjournal.com/article/CA6626579.html?nid=3603 (accessed Oct. 31, 2010).

Notes on Operations Digital Curation Planning at Michigan State University

Lisa Schmidt, Cynthia Ghering, and Shawn Nicholson

Recognizing the need for guiding the management and preservation of Michigan State University's digital assets, a team led by university archivists and librarians conducted a digital curation planning project to explore and evaluate existing digital content and curation practices. The team used data gathered in this study to identify next steps in digital curation planning, including the recommendation to collaborate with other universities to develop solutions. While the findings were specific to Michigan State University, the process of assessing practices and identifying needs may be replicated elsewhere.

Research universities and other large organizations continue to create and amass large volumes of digital assets and information. The inherent fragility of digital material because of technology obsolescence and physical threats such as media instability put these valuable digital assets at risk of eventual inaccessibility. In 2009, Michigan State University (MSU) confronted the problem of management and long-term preservation (curation) of its digital assets by implementing a digital curation planning project. The University Archives and Historical Collections (UAHC), MSU Libraries, and MATRIX: Center for Humane Arts, Letters, and Social Sciences Online collaborated in this exploration of digital asset management at the university with the goal of developing campuswide guidelines for good practices in digital curation.

This paper describes the MSU project beginning with the institutional context and challenges of digital asset management faced by a major research university and explains the plans, methods, and results of the study. The project team developed and implemented a self-selective, campuswide survey of digital assets and technological infrastructures using a web-based questionnaire. After the project team analyzed the results of the questionnaire, they selected eleven units for in-depth, one-on-one interviews regarding their digital curation practices. The paper concludes with recommended next steps in digital curation planning for MSU. The MSU study may serve as a model for similar research at other universities; the MSU recommendations also may be relevant.

Throughout this paper, the term "digital" is defined as "representing information through a sequence of discrete units, especially binary code."[1] The terms "digital content," "digital resources," "digital data," "digital assets," "digital material," "digital objects," and "digital information" are essentially interchangeable. "Digital media" refers to the physical electronic media that holds digital information, such as magnetic tape, CDs, and computers used as file servers. "Digital asset management" and "digital curation practices" refer to the handling of digital material. "Digital preservation" is defined by the California Digital Library as "the managed activities necessary for ensuring the long-term retention and usability of digital objects."[2] "Digital curation" includes preservation but also takes into

Lisa Schmidt (lschmidt@ais.msu.edu) is Electronic Records Archivist, University Archives and Historical Collections, Michigan State University; **Cynthia Ghering** (ghering@ais.msu.edu) is Director, University Archives and Historical Collection, Michigan State University; and **Shawn Nicolson** (nicho147@mail.lib.msu.edu) is Assistant Director for Digital Information, Michigan State University Libraries, East Lansing, Michigan.

Submitted July 6, 2010; returned to authors August 8, 2010, with request to revision and submit; revision submitted October 6, 2010, and reviewed; tentatively accepted November 21, 2010, pending modest revision; revision submitted Dec. 20, 2010, and accepted for publication.

account the life cycle of the data to include its creation and management. As defined by the Digital Curation Centre, "Digital curation is maintaining and adding value to a trusted body of digital information for current and future use . . . the active management and appraisal of data over the life-cycle of scholarly and scientific materials."[3]

Literature Review

Cultural institutions have been concerned about the increasing number of digital objects for more than two decades. Writing to the museum community in 1994, Zorich warned of the pervasive problems associated with "converting their paper-based records into digital form" resulting in "millions of records."[4] She drew upon the nascent experience of the U.S. Department of Energy (DOE), as that government agency proposed the then novel idea that electronic data created from the Human Genome project be "curated." She extended the notion of digital curation first used by the U.S. DOE so that it could be applied to the digital documentation of museum collections. Focusing on the need for "consistency, long-term quality, and relevance over time," and pointing out that any changes require verification and authentication, Zorich presciently described challenges that institutions continue to face today.[5]

Research universities are creating and storing digital information in volumes unthinkable to early writers such as Zorich. In the 2009 inaugural issue of the *International Journal of Digital Curation*, Tindermans described the concerns of the digital curation community, observing that "the sheer volume of digital data, the bewildering variety of formats and digital objects, which sometimes turned up in fast-changing sequences of new versions, soon became a cause of great concern."[6] Without active, well-considered, long-term plans for managing and preserving these resources, these digital assets eventually will become inaccessible because of technology obsolescence and digital media fragility. In a 2004 article describing electronic records strategies for small institutions, Cook stated, "once these digital records have been created or captured, they must then be preserved . . . each has an optimistic shelf life in digital format of perhaps 20 years before significant archival intervention is needed . . . before it either disappears as unreadable or self-destructs physically."[7] Many organizations also face digital storage limitations because of the increasing size of video and other multimedia files, as well as the increasing volumes of files in general. Although storage costs have decreased, budgets have tightened in the current economy. Litigation liabilities and search difficulties result when increasing volumes of digital assets are created and stored indefinitely.

In the late 1990s and into the twenty-first century, the digital curation community produced several guidelines and best practices documents for digital preservation and curation. Among them are the Digital Curation Centre's "Curation Reference Manual," the Digital Preservation Coalition's "Digital Preservation Handbook," and the United Nations Educational, Scientific, and Cultural Organization's (UNESCO) *Guidelines for the Preservation of Digital Heritage*.[8] These handbook-type materials represent an important stage in the evolution of the community. They bring focused attention to practices familiar to archivists and librarians, such as appraisal, selection, metadata creation, storage, security, and future planning, and place them into the digital context. The handbooks are also important because they contain advice and provide best practice for areas not often as comfortably familiar to librarians and archivists, including selecting file formats, using open-source software, performing bit-level integrity checks, and employing migration and emulation techniques.

As the evolution continued, the community created two critical documents that developed a broad framework to understand what is required to provide long-term stewardship of digital information. The Center for Research Libraries (CRL) and Online Computer Library Center (OCLC) issued *Trusted Repositories Audit and Certification (TRAC): Criteria and Checklist*, based on the International Standards Organization's Open Archival Information System (ISO OAIS) reference model, which approaches the problem of good digital preservation practice by articulating eighty-four criteria for a trusted digital repository.[9] The criteria are organized into three sections: "Organizational Infrastructure," which includes administrative, management, and financial support, as well as copyright; "Digital Object Management," which includes the accession, storage, and provision of access to digital objects; and "Technologies, Technical Infrastructure, and Security," which includes backup and disaster recovering planning. All criteria require evidence of compliance, preferably as written documentation and policies.

Research data poses a different set of challenges and its own growing body of literature. As it relates to this project, Gold's review of the library's role in data curation highlighted milestones such as the establishment of a data curation education program at the University of Illinois and the Johns Hopkins Sheridan Library's leadership of a multi-institutional effort to build a national infrastructure for curation of research datasets.[10] Offering a more granular view, Walters provided a retelling of Georgia Institute of Technology's journey toward curating scientific data and developed a roadmap that can easily be adapted to other institutions.[11] Mirroring the broader community's evolution,

Abrams, Cruse, and Kuntz emphasized the need to constantly reevaluate and refocus curatorial activities.[12] They discussed how the California Digital Library has moved away from a singular focus on preservation toward a more encompassing "programmatic, rather than a project-oriented approach; and a renewed emphasis on services, rather than systems."[13]

As new tools and expectations arise, richer exploration options and an expansion of existing best practice are needed. The results of two large-scale surveys that explored perceived needs and current practices were published in 2009. Sinclair and colleagues surveyed 172 national libraries, archives, and other cultural heritage organizations in Europe to "better understand the organization's digital preservation activities and needs."[14] While the respondents expressed a high level of awareness of the challenges of preservation, the authors found no corresponding level of policy and budgetary development. An Educause Center for Applied Research study by Yanosky examined data management in U.S. higher education.[15] Yanosky employed a multifaceted research design consisting of a quantitative web-based survey, qualitative interviews, and case studies; 309 responded to the survey portion and 23 participated in targeted interviews. The results confirmed the vast amount of data (described as "data big bang") that is being created and that data stewardship and security are significant concerns. When asked about enterprise-wide content management, "only about 12 percent of 304 respondents [said that their institution] has an integrated enterprise content management solution."[16]

Although the literature is replete with discussions of digital curation theory development and evolving best practices, no known studies describe an institution-wide effort to understand curatorial activities across the full range of digital objects.

Institutional Context and Challenges of Digital Asset Management

Like other research universities, MSU has amassed a growing body of digital assets and information—including institutional records, faculty and student research, theses and dissertations, university publications, multimedia collections, digital surrogates of cultural material, learning objects, course materials, and more. Much time, effort, grant funding, human capital, and research have gone into creating these digital resources—some of which only exist in digital form. In response to the need to manage digital content, some MSU colleges and departments have started their own digital repositories. No comprehensive, campuswide digital preservation strategy or set of guidelines exists, however, and MSU has not established an institutional repository.

The University Archives and Historical Collections (UAHC) (http://archives.msu.edu/index.php), the official repository for MSU's historical archives, began looking into the problem of cross-campus digital preservation in early 2009. Along with its mandate to collect, preserve, and provide access to the historical records of the university, the UAHC also is charged with assisting departmental and administrative units in the efficient administration, management, and preservation of all official university records, including electronic records.

In the winter of 2009, the UAHC engaged a digital preservation intern from the University of Michigan's School of Information to research a strategy for preserving and making accessible MSU's multimedia productions, including born-digital film, audio, and still images. This internship was funded through an Institute of Museum and Library Services (IMLS) grant.[17] Under the direction of the UAHC director, the intern identified file formats, recommended workflows, and addressed the long-term preservation of electronic media. During the course of the project, UAHC staff and the intern held interviews with stakeholders at seven campus units: the Digital Media Center/Vincent Voice Library at MSU Libraries; MATRIX; Michigan Government Television, a public affairs initiative of Michigan's cable television industry; Sports Broadcasting; University Relations, which handles communications and public relations; Virtual University Design and Technology, which specializes in creating online courses and e-learning tools; and Broadcasting Services, including the WKAR public television and radio affiliates.

The project concluded with a report to the UAHC director that included several high-level recommendations for the creation, storage, preservation, and long-term access to multimedia university records. The intern suggested surveying more campus units to broaden the analysis of the university's digital assets, creating best practice guidelines for selecting and appraising content, file formats, naming conventions, and metadata; providing better long-term storage options; and establishing an institutional repository. This digital preservation internship project, while not groundbreaking, raised staff awareness and led to a more formalized and significant environmental scan of the institution's strengths and weaknesses in this area.

Building on this initial research, the director proposed developing a more comprehensive environmental scan and gap analysis inspired by the Inter-University Consortium for Political and Social Research (ICPSR) digital preservation management workshop.[18] The purpose of the project was to develop a digital preservation plan with a focus on practical solutions using MSU's current resources; the initiative recommended a collaborative effort of the UAHC, the MSU Libraries, and MATRIX.[19] The

resulting proposal received top-level support, with the university's vice provost of libraries, computing, and technology committing funding for a half-time digital preservation analyst to manage the project for one year beginning in July 2009.

To that end, the director of UAHC created a cross-university Digital Preservation Planning Project team and brought in an electronic records archivist then working for MATRIX to serve as the digital preservation analyst. The team, which served to guide the analyst's work and to encourage participation from institutional stakeholders, consisted of staff from UAHC, the MSU Libraries, MATRIX, MSU Extension/Agriculture and Natural Resources Technology Services, campus Information Technology's (IT's) data services, and the Office of the Registrar. During monthly meetings, the team reviewed activities and progress, anticipated next steps and potential challenges, and analyzed findings.

As the project got underway, the team quickly realized that the initial plan of an all-campus environmental scan was overly ambitious. The team could not complete an exhaustive survey of all university digital assets within one year. Even if more staff resources were available, an all-encompassing inventory would require a significant effort and result in diminishing returns because of inevitable redundancies. Perhaps most important, the team was concerned that a comprehensive inventory would create the perception that the project's desired outcome was to build a one-size-fits-all data repository. In MSU's heavily decentralized environment, both academic and administrative units are apprehensive of solutions that may result in loss of control of the digital assets of their college or unit. Therefore, instead of conducting a comprehensive survey, the team decided to focus on the variety of digital content and the disparate needs of the different campus units.

To address the concerns stated above, the project team refined the scope of work and deliverables. The team decided to change the name of the project, replacing the term "digital preservation" with "digital curation." This new project name emphasized the focus of the study on the life cycle of the university's digital assets, including preservation, creation, and management. Instead of conducting a comprehensive inventory, the team administered an online survey that would provide a sampling of digital assets and repositories based on a campus-wide, self-selective, web-based questionnaire. The team later conducted in-depth interviews with a limited number of units, with consideration given to learning more about the digital repository solutions already implemented across campus. Technical infrastructures, storage needs, metadata schemes, and file naming conventions also were reviewed.

By offering practical guidance and influencing policy, MSU's archivists and librarians sought to expand on their traditional role as custodians of physical material. As Cox stated in a 2008 *EDUCAUSE Review* article, "The institutional archive needs to assume more of a policy role, identifying records throughout the campus and working to ensure that digital records are both maintained by their creators and kept ready for research use."[20] The information professional as policymaker is not a new idea. Bearman maintained in a 1991 paper that in the interests of gaining respect as professionals, archivists and records managers should move away from the custodial role and "reposition themselves as policy makers and regulators whose task it is to assure that managers throughout the organization demonstrate awareness of the institutional significance of information by retaining and destroying information at appropriate times and in appropriate ways."[21] Thus MSU's archivists and librarians can provide guidelines and best practices in digital preservation and management while the material itself remains in the custody of the creating units.

Self-Selective, Campuswide Survey

Research Method

In October 2009, the digital curation planning project team designed and conducted a voluntary, web-based survey using a questionnaire created with the online SurveyMonkey tool. UAHC publicized the survey through MSU's internal e-mail–based communications network for information technology employees, the weekly *MSU News* online newsletter, and the project website (http://msudcp.archives.msu.edu). These notifications emphasized the need for digital curation planning and encouraged involvement of technology staff and content creators in the survey as a first step. At the same time, the strictly voluntary nature of participation was stressed. The digital curation planning project team made the questionnaire available for two weeks.

Survey participants were asked about types of digital content created by their units, the approximate volume of digital content in terabytes, storage media used, file formats created, online storage capacity and storage expansion plans, and any content management system (CMS) or digital repository software used. (Neither "CMS" nor "digital repository software" was defined in the questionnaire, allowing survey takers to respond according to their own interpretations of the terms.) The questionnaire appears in appendix A.

Findings

The questionnaire received 90 responses, including 23 responses

from academic departments, 31 responses from administrative services units, 9 responses from research centers and institutes, and 27 responses from technology services organizations. This is a good response rate for a large research university such as MSU. As of fall 2009, MSU offered more than 200 programs from 17 academic degree–granting colleges, had a student body of more than 47,000, and supported more than 50 administrative services units and 70 research centers.[22] Responses that came from organizations that provide technical services were placed in the "technology services" category, even if those organizations officially belonged to academic or administrative units. Because the link to the questionnaire was distributed through an e-mail list and made publicly available through the online MSU staff newsletter, multiple staff per unit could respond.

Academic departments represented in the questionnaire responses covered a wide range of fields, from agricultural economics, nursing, and veterinary medicine to math and science education, physics and astronomy, telecommunications, business, athletics, and the arts. Participating administrative units ranged from the Controller's Office, the Capital Asset Management Department, the Office of Planning and Budget, the Office of the President/Board of Trustees, and the MSU Libraries to Broadcasting Services, University Relations, and Virtual University Design and Technology. The research centers included the Confucius Institute, the Cyclotron, the Julian Samora Research Institute, and MATRIX. In contrast, all of the technology services responses came from only 6 organizations.

The types of digital content making up the largest proportion of a given unit's content varied considerably. Digital and scanned photos and images, word processing documents, and research data sets topped several of the academic departments' lists while administrative units reported large proportions of imaged or scanned paper documents, word processing and spreadsheet documents, and databases. Research data, audiovisual material, word processing documents, and programming code predominated at the research centers. The technology services organizations noted that most of their digital content consisted of code, databases, and webpages.

File formats comprising the largest proportion of a given unit's digital content also varied. The academic departments surveyed noted the existence of PDFs, Statistical Package for the Social Sciences (SPSS) and Statistical Analysis System (SAS) statistical formats, TIFFs, JPEGs, MySQL, and Camtasia video formats. Various database formats, TIFFs, text, MS Office formats, as well as audio and video formats, prevailed at the administrative units. The research centers reported sizeable concentrations of video formats, PHP code, MS Word, and SAS, and the technology services organizations carried large proportions of text and programming code formats.

Most of the units reported that they store digital content on hard drives and most also use some combination of different types of removable media as well as network storage; one unit reported storing data on cassette tapes. Seventeen units plan to increase online storage capacity soon, most from 1 to 10 terabytes, with 6 units planning to add 30 terabytes or more.

Twenty-three units responded that they have implemented or plan to implement a CMS, digital repository software, or both. Again, neither "CMS" nor "digital repository" was defined in the questionnaire, so unit representatives responded using their understanding of the terms. CMSs noted include Sharepoint, Alfresco, Mura CMS, Drupal, Cascade CMS, Document Viewer, and DotNetNuke, as well as in-house-developed solutions; one unit reported using Trac Project, an issue-tracking system for software development projects. The Physical Plant Division uses the Facilities Administration Management Information System (FAMIS) to manage operations, maintenance, and repair projects for the university's physical environment. Digital repository solutions included KORA, the Madison Digital Image Database (MDID), ResourceSpace, Portfolio Server 9, and MSU Extension's custom-built Knowledge Repository system. In some cases, the same software solution was listed as the CMS and the digital repository application. Tools (including Concurrent Versions System (CVS), Git, Adobe Version Cue, and Subversion), more properly known as version control software, were identified as CMSs or digital repository software by some units. The database application FileMaker was listed as a CMS by one unit. Some units did not use a software tool, but instead noted that they managed content on web and file servers, or that they used homegrown solutions such as wikis.

Many respondents provided additional comments stating great interest and enthusiasm in the digital curation planning project's goal of providing curation guidance. One administrative unit noted, "This is a timely survey, because our unit is at a point where we have to choose which data to delete off our servers, as we are accumulating more than we can afford to store. We need university guidelines and related archival resources." Another unit asked for guidelines on how to handle archive-worthy files at the time of creation rather than storing everything and later subjecting the unit to an arduous appraisal process. Respondents also expressed interest in guidance on choosing a digital asset management system.

One-on-One Interviews

Research Method

The digital curation planning project team chose 11 units that responded to the survey to interview, focusing on the units that reported using a CMS, digital repository solution, or both. The team determined that these units would have concentrations of digital content they intentionally manage or preserve and thus would be good resources for the project. These units might already have solutions in place that could be folded into guidelines useful to other units. Conversely, these units might need the most help and guidance from the project. The team also was interested in talking to offices that created digital content documenting MSU history or generated university records that fell under existing institutional records retention schedules. Units interviewed were

- Broadcasting Services (www.wkar.org), including the WKAR public radio and TV stations, which produces digital content of historical value to the university that should be transferred to UAHC.
- Center for Research on Mathematics and Science Education (CRMSE), which focuses on grant-funded research into the teaching of mathematics and science. CRMSE is primarily concerned with the preservation of survey instruments and research data, and the unit is interested in transferring these files to UAHC for long-term preservation.
- Confucius Institute at Michigan State University (CIMSU) (www.experiencechinese.com), which is part of a worldwide network that promotes the teaching of Chinese language and culture. The unit creates multimedia content for use in online training and educational games.
- Department of Art and Art History (www.art.msu.edu), which hosts the Visual Resources Library (VRL) of digital art images.
- Department of Theatre (www.theatre.msu.edu), which stores a mix of digital photos of performances, CAD drawings of sets, and other performance-related digital files.
- MATRIX: Center for Humane Arts, Letters, and Social Sciences Online (www.matrix.msu.edu), a digital humanities research center and host for major digital libraries of cultural material, including the African Online Digital Library (AODL), Detroit Public Television's *American Black Journal* video archives, Historical Voices, and the Quilt Index.
- MSU Extension/Agriculture and Natural Resources (ANR) Technology Services, which provides communities across Michigan with programming focused on agriculture and natural resources; children, youth, and families; and community and economic development.
- National Superconducting Cyclotron Laboratory (NSCL) (www.nscl.msu.edu), a National Science Foundation (NSF)–funded, world-leading laboratory for rare isotope research and nuclear science education.
- Physical Plant Division (www.pp.msu.edu), which maintains all buildings and land entities on and off campus and provides utility services. In addition to its operations-related digital management systems, Physical Plant holds architectural drawings and building maps that may have historical significance.
- Turfgrass Information Center (TIC), MSU Libraries (http://tic.msu.edu), which manages the Turfgrass Information File (TGIF) database of published and unpublished materials relating to turfgrass science, culture, and the management of turfgrass-based facilities, such as golf courses, parks, sports fields, lawns, sod farms, roadsides, institutional grounds, and other managed landscapes.
- University Relations (http://www2.ur.msu.edu), the unit responsible for public relations, creates and maintains large volumes of digital photographic and video records of archival value, some of which should be transferred to UAHC. Digital video footage includes the *MSU Today* show on the Big Ten Network (www.bigtennetwork.com), a joint venture between Fox cable television networks and the Big Ten Conference, a U.S. collegiate athletics organization.

Digital curation planning team members conducted the one-on-one interviews January–March 2010. The meetings were informal, two-hour conversations held at the unit offices. Discussion topics included how the unit's digital content relates to the mission of the unit, whether it was of ongoing use or of archival value—that is, whether it documents the activity of the unit or the university file formats used, and storage infrastructure, including any space issues.

The team asked about the CMS(s) or digital repository software used, why they were chosen, and how they are used. The team also asked about processes and workflows of ingesting data and digital content into the system, archival storage and preservation of content, and content retrieval, as

well as whether the unit had a means to ensure file integrity. Finally, the team asked about the use of metadata stored with or related to the content as well as file naming conventions. (See appendix B for the types of questions asked.)

Findings and Analysis

In analyzing the results of the interviews, the team noted several key points. Each unit had devised solutions that fit the mission of the unit, the nature of the data, and the needs of users. Some units use commercial applications and some use open-source software. The Turfgrass Information Center at the MSU Libraries, for example, has long used the commercially available Cuadra STAR database and CMS, and the Department of Theatre uses the relatively new open-source ResourceSpace digital repository solution. See appendix C for a list of the content management, digital repository, and other software used by the interviewed units to manage digital assets.

Some units, such as Broadcasting Services, hold digital content of archival value to the university. Other units, such as the Department of Art and Art History and MATRIX, create and store digital materials that is produced on behalf of partner cultural organizations. These are both important digital resources, but are not institutional records that must adhere to a defined retention schedule.

Many of the interviewed units exhibited good digital preservation and curation practices. Most backed up their data in some manner. Some used detailed metadata to describe their digital assets, and many were using repository software with good access and discovery interfaces to manage their content. Many of the units have strong support from their administrative management and stable funding.

Most of the interviewed units store preservation (archival) masters of at least some of their content. MATRIX maintains TIFF files of images and preservation masters of audiovisual content that has been converted to access formats for use in the KORA digital repository. The Turfgrass Information Center stores scans of printed material and slides as TIFF files but makes them available online in PDF and JPEG formats. Likewise, the Department of Art and Art History keeps TIFF master files while providing JPEG files as access copies in the Visual Resources Library. The Department of Theatre, on the other hand, has chosen to convert digital photos from the original RAW format to JPEGs for use in its DOT::Media repository. Preservation masters of MSU Extension's bulletins are stored as TIFF files in a dark archive at the MSU Libraries, with PDF versions available through the MSU Extension Knowledge Repository. Both Broadcasting Services and the Confucius Institute maintain some audio files in the archival WAV format. Physical Plant currently stores and scans documents in TIFF but would like to move to PDF/A as a preservation master format. The Center for Research on Mathematics and Science Education (CRMSE) wishes to convert data sets to XML and survey instruments to PDF/A files for long-term preservation.

Only three units shared their means for verifying file integrity. MATRIX's KORA repository software includes a message digest algorithm that can generate a unique number for an ingested file and then periodically check that number; any change would indicate that the file had become corrupt. Using Adobe Bridge, the Department of Art and Art History can detect file corruption when viewing thumbnail photos; likewise, the Adobe Photoshop script that creates contact sheets of thumbnails will stop running if it encounters a corrupt file. Soon the Department of Theatre will add a file integrity test to the code for the DOT::Media repository, accompanied by the capability to store parity files, which record the structure of files that need to be protected. If the original files become corrupted, the parity files may be used to restore them.

MSU Extension and the Department of Art and Art History have formal file naming conventions in place. The Physical Plant Division's Meridian system automatically assigns file names that include the project number, document type, and a brief, metadata-based code. Although the Cyclotron has a systematic method of assigning project numbers to file directories for each experiment, researchers have some latitude in naming the actual data files. MATRIX develops file-naming conventions with its partners on a project-by-project basis.

Most of those interviewed expressed interest in curation guidelines and said they could use guidance. Although these units back up their data, most of the backups tend to be located very close to production servers—often in the same building, if not the same room. The high incidence of maintenance of preservation copies is encouraging, but not practiced by all units and for all file types. Alternatively, the practice of checking file integrity is disappointingly low. Some of the units create only minimal metadata for their digital content, and the project team found little in the way of documented digital curation policies. The lack of good digital curation practice is unfortunate considering that the interviewed units are more likely to have invested in digital asset management compared to the rest of campus. The team suspects that most campus units are either unaware of the need for or unable to address digital curation at this time.

The team also compared the metadata schema used by the units interviewed with the Dublin Core (DC) metadata element set, a standard in the information science field for describing digital objects.[23] The comparison involved examining the metadata element sets used by the units, noting correspondences with the DC element set, and reviewing definitions of elements in data dictionaries for schema when necessary to understand the information that they were intended to represent. For example, MSU Extension's "Author One" and "Author Two" elements correspond to the DC "Creator" element. In some cases, the analyst contacted original unit interviewees for in-depth explanations of particular elements.

Six of the units interviewed had implemented metadata schema that could be considered in the comparison. MSU Extension, MATRIX, and the Department of Theatre use metadata schema based on or similar to DC, with slight variations to reflect local needs. The Department of Art and Art History uses the Image Resource Information System (IRIS) cataloging utility for describing its art images with metadata based on the Visual Resources Association (VRA) Core and the Cataloging Cultural Objects (CCO) guide to good cataloging practices; this metadata maps roughly to DC.[24] The metadata used by Physical Plant and the Turfgrass Information Center, the other two units in the comparison, do not correspond directly to the DC metadata set. Physical Plant uses the metadata customization capabilities of the Meridian facilities assets management system to specify its own locally controlled vocabulary suited to project transactions. Likewise, the Turfgrass Information Center uses its own indexing terms specified in the Cuadra STAR system for creating descriptive metadata for bibliographic information about physical and digital material related to turfgrass.

Digital Storage Solution Planning at Michigan State University

In tandem with the exploration into digital curation practices, the digital curation planning project team kept abreast of and helped to influence digital storage planning at MSU. The Libraries, Computing, and Technology (LCT) division provides central technology support for administrative business systems, e-mail, academic, and network services throughout the campus. A strong tradition of local units maintaining their own IT staff and managing their own systems also exists, however. Like many other institutions, MSU is looking closely at this divide between centralized and local IT to discern how best to use and consolidate storage and other technology functions. One recognized advantage of a strong central IT organization is that it can more effectively manage electronic records and digital assets.

Currently, LCT is developing virtual server environments and price structures as a storage solution to local units. LCT also is considering tiered storage options, which entail a variety of storage types or levels to meet a diverse array of needs. The storage solution would be tiered depending on the nature of the content. For example, storage space for files of temporary, short-term use might be provided locally while capacity and infrastructure efficiencies would be leveraged by developing a centralized, shared long-term preservation storage environment.

The Committee on Institutional Cooperation (CIC), a consortium of Big Ten universities and the University of Chicago, is developing collaborative storage solutions that would better enable effective, efficient stewardship of campus assets and cutting-edge scholarship. The CIC expects common architecture, infrastructure, and operating environments to increase economies of scale, permit shared management, and provide for the research and development of such value-added services as community tagging and annotation, citation tracking, and digital curation for scholarly data. Many of these efforts derive from the CIC chief information officers' 2010 report, "A Research Cyberinfrastructure Strategy for the CIC: Advice to the Provosts from the Chief Information Officers."[25] These collaborative storage solutions are separate from HathiTrust (www.hathitrust.org), which aims to build a reliable and comprehensive digital archive of library materials converted from print that is co-owned and managed by a number of academic institutions and in which the CIC partners.

Next Steps in Digital Curation Planning

The digital curation planning project team observed that despite the variety of digital content and formats as well as different approaches to curation, commonalities and patterns between the units interviewed emerged. Studying these patterns led to the identification of four basic types of digital content created at MSU: university publications, including e-journals and electronic theses and dissertations; digital content that documents the history of the university, such as the photos and video of University Relations and some of Broadcasting Services' audiovisual programming; nonuniversity digital content, such as the digital files created and managed by MATRIX and the Department of Art and Art History; and research data.

With an understanding of the types of digital content in hand, the curation needs of the units can be addressed and functional specifications for curation solutions developed and applied as needed. For example, university publications will require deposit in an institutional repository

and the implementation of a distributed preservation solution such as Lots of Copies Keep Stuff Safe (LOCKSS) (www.lockss.org). Digital content that documents the history of the university will require digital curation and appraisal guidelines as well as mechanisms for transfer to and storage with UAHC. Digital content not specific to MSU will benefit from curation guidelines on metadata, file formats, repository software, file integrity checking, consistent file naming conventions, and storage and backup planning. Units that create and manage research data will require support in meeting the new National Science Foundation (NSF) requirements for grant proposals to include a data management plan that addresses preservation, access, and other elements of digital curation.[26]

In keeping with the approach of identifying user needs and developing functional requirements, the digital curation planning project team recommended several next steps to guide the management and preservation of MSU's digital assets. For example, UAHC can provide digital asset appraisal assistance to university departments and units, especially those holding digital assets of historical value that should be transferred to UAHC for long-term preservation, as is planned for University Relations. Work will continue with LCT on the development of tiered storage plans as well as plans for transferring digital content of archival value to UAHC. In that same vein, UAHC will develop guidelines to quickly determine whether digital assets should be transferred for permanent preservation.

UAHC, the MSU Libraries, and representatives from the other members of the digital curation planning project team will work together to explore the digital curation practices of units holding significant digital content that were not represented in this project, particularly those with other types of content and with different content management practices; this will include further investigation of research data curation across the campus. They will develop general best and good practices in digital curation recommendations and guidance, keeping in mind differences in the missions of the units and the types of digital material that they create and manage.

With the MSU Libraries and other digital curation planning team principals, UAHC will develop metadata standards for university records, including publications and digital content that documents the history and activity of the university. UAHC and the MSU Libraries will work together to develop digital data curation toolkits that acknowledge researchers and units as information producers as well as consumers. Topics covered by these toolkits will include file formats, documentation, intellectual property rights, sharing and dissemination, and preservation.

The team also recommended the fostering of "Communities of Practice" through online forums and meetings, in which campus units and other institutions have the opportunity to share digital curation experiences, generate new ideas, and collaborate on initiatives. UAHC and the MSU Libraries could work with their counterparts in the digital humanities and at other CIC member institutions to obtain grant funding to explore the digital curation problem across institutions and develop common best and good practices guidelines.

Finally, the team advocated the creation of a senior-level digital preservation officer position at MSU. This individual could continue to raise the visibility of digital curation, focus the coordination of curation and preservation resources across campus for both academic and administrative data types, and direct the dissemination of digital curation guidelines and best practices. Although economic conditions prohibit hiring a senior-level digital preservation officer for the university now, a new digital curation librarian position has been created in the Digital Information Services unit of the MSU Libraries.

Conclusion

In response to the need for ensuring the viability of the valuable digital assets created in ever-larger volumes at MSU, a year-long digital curation planning project explored current practices in the creation and management of digital material. Data was gathered and observations made through a self-selective, campuswide online survey and one-on-one interviews with campus units that demonstrated some level of curation practice or held material of historical interest to the university.

Although the digital curation planning project team initially found the variety of digital content and formats overwhelming, patterns emerged that will make addressing the problems of digital curation at MSU easier. Four types of digital content were identified, and user needs can be articulated and functional specifications developed to meet those needs. By investigating the needs of more campus units and continuing to build on these patterns, digital curators can make sense of the jumble of digital content at MSU and develop solutions that also may be of use to other institutions.

Developing digital curation guidelines for Michigan State University will be an iterative process. Recommended next steps include UAHC providing appraisal assistance to units holding material of archival value, such as University Relations; studying and influencing the curation of scholarly research data; developing of common metadata standards and curation toolkits; fostering of "Communities of Practice" within the university and with partner institutions; and working with other universities to obtain grant funding for the study of digital curation practices across institutions. By

surveying and interviewing a variety of departmental and administrative units, the digital curation planning project team began developing an understanding of the digital assets and related needs and concerns of the MSU community, building trust, and establishing new relationships that will aid in moving forward with an institutional approach to digital curation planning. While the MSU findings are specific to one institution, the process of assessing practices and identifying needs can be replicated elsewhere.

References

1. Richard Pearce Moses, "A Glossary of Archival and Records Terminology," Society of American Archivists, www.archivists.org/glossary/term_details.asp?DefinitionKey=340 (accessed Nov. 30, 2010).
2. California Digital Library (CDL), "Glossary," http://www.cdlib.org/gateways/technology/glossary.html?field=institution&query=CDL&action=search, (accessed Nov. 30, 2010).
3. Digital Curation Centre, "DCC Charter and Statement of Principles," 2010, www.dcc.ac.uk/about-us/dcc-charter (accessed Apr. 9, 2010).
4. Diane Zorich, "Data Management: Managing Electronic Information: Data Curation in Museums," *Museum Management & Curatorship* 14, no. 4 (1995): 431.
5. Ibid., 430–32.
6. Peter Tindermans, "Key Stakeholders Pledge to a Strategic Approach to Preserve the Digital Records of Science," *International Journal of Digital Curation* 1, no. 1 (2009): 84, www.ijdc.net/index.php/ijdc/article/viewFile/13/8 (accessed Dec. 23, 2010).
7. Terry Cook, "Byte-ing Off What You Can Chew: Electronic Records Strategies for Small Archival Institutions," *Archifacts* (Apr. 2004), www.aranz.org.nz/Site/publications/papers_online/terry_cook_paper.aspx (accessed May 26, 2010).
8. Digital Curation Centre, "Curation Reference Manual," 2010, www.dcc.ac.uk/resources/curation-reference-manual (accessed June 21, 2010); Digital Preservation Coalition, "Digital Preservation Handbook," (2009), www.dpconline.org/advice/preservationhandbook (accessed Sept. 20, 2010); National Library of Australia, *Guidelines for the Preservation of Digital Heritage* (United Nations Educational, Scientific and Cultural Organization, 2003), http://unesdoc.unesco.org/images/0013/001300/130071e.pdf (accessed June 21, 2010).
9. The Center for Research Libraries and Online Computer Library Center, *Trustworthy Repositories Audit and Certification: Criteria and Checklist*, version 1.0 (Chicago: CRL; Dublin, Ohio: OCLC, 2007), www.crl.edu/sites/default/files/attachments/pages/trac_0.pdf (accessed June 2, 2010); Consultative Committee for Space Data Systems, *Reference Model for an Open Archival Information System (OAIS)*, CCSDS 650.0-B-1 Blue Book issue 1 (Washington, D.C.: CCSDS, 2002), http://public.ccsds.org/publications/archive/650x0b1.pdf (accessed Sept. 13, 2010).
10. Anna Gold, "Data Curation and Libraries: Short-Term Developments, Long-Term Prospects" (Apr. 14, 2010), http://digitalcommons.calpoly.edu/lib_dean/27 (accessed Sept. 29, 2010).
11. Tyler Walters, "Data Curation Program Development in U.S. Universities: The Georgia Institute of Technology Example," *International Journal of Digital Curation* 3, no. 4 (2009): 83–92, www.ijdc.net/index.php/ijdc/article/viewFile/136/153 (accessed Dec. 23, 2010).
12. Stephen Abrams, Patricia Cruse, and John Kunze, "Preservation Is Not a Place," *International Journal of Digital Curation* 4, no. 1 (2009): 8–21, www.ijdc.net/index.php/ijdc/article/viewFile/98/73 (accessed Dec. 23, 2010).
13. Ibid., 8.
14. Pauline Sinclair et al., "Are You Ready? Assessing Whether Organisations are Prepared for Digital Preservation," (paper presented at iPRES 2009: The Sixth International Conference on Preservation of Digital Objects, Oct. 5–6, 2009), www.escholarship.org/uc/item/8dd2m5qw (accessed Sept. 13, 2010).
15. Ronald Yanosky, *Institutional Data Management in Higher Education*, Research Study no. 8 (Boulder, Colo.: EDUCAUSE Center for Applied Research, 2009), www.educause.edu/Resources/nstitutionalDataManagementinH/191754 (accessed Sept. 29, 2010).
16. Ibid., 17.
17. University of Michigan, Engaging Communities—Fostering Internships for Preservation and Digital Curation, "About the Project," http://preservation.cms.si.umich.edu (accessed Dec. 23, 2010).
18. Anne R. Kenney et al., "Digital Preservation Management: Implementing Short-Term Strategies for Long-Term Problems," (online tutorial and workshop, Cornell University Library, 2003, version 1); Inter-University Consortium for Political and Social Research (ICPSR), University of Michigan, 2007), www.icpsr.umich.edu/dpm/dpm-eng/eng_index.html (accessed May 25, 2010).
19. MSU Digital Preservation Proposal—April 2009: Project: Preserving MSU's Digital Assets, http://msudcp.archives.msu.edu/wp-content/uploads/2010/12/MSU-DPP-Proposal.pdf (accessed Dec. 23, 2010.)
20. Richard Cox, "The Academic Archives of the Future," *EDUCAUSE Review* 43, no. 2 (2008), www.educause.edu/EDUCAUSE+Review/EDUAUSEReviewMagazineVolume43/heAcademicArchivesoftheFuture/162683 (accessed June 3, 2010).
21. David Bearman, "An Indefensible Bastion: Archives as a Repository in the Electronic Age," *Archival Management of Electronic Records*, *Archives and Museum Informatics Technical Report* 13 (Pittsburgh: Archives & Museum Informatics, 1991): 17.
22. Michigan State University, "MSU Facts," 2010, www.msu.edu/about/thisismsu/facts.html (accessed Oct. 6, 2010).
23. Dublin Core Metadata Initiative, "Dublin Core Metadata Element Set, Version 1.1," Oct. 11, 2010, http://

dublincore.org/documents/dces (accessed Dec. 23, 2010).
24. Visual Resources Association, "VRA Core Four," www.vraweb.org/projects/vracore4/ (accessed May 14, 2010); CCO Commons: Cataloging Cultural Objects, "Cataloging Cultural Objects: A Guide to Describing Cultural Works and Their Images," www.vraweb.org/ccoweb/cco/about.html (accessed May 24, 2010).
25. Committee on Institutional Cooperation Chief Information Officers, "A Research Cyberinfrastructure Strategy for the CIC: Advice to the Provosts from the Chief Information Officers," 2010, www.cic.net/Libraries/Technology/2010Report_-_6_21reduced.sflb.ashx (accessed Sept. 16, 2010).
26. National Science Foundation, "Scientists Seeking NSF Funding Will Soon Be Required to Submit Data Management Plans," Press Release 10-077, May 10, 2010, www.nsf.gov/news/news_summ.jsp?cntn_id=116928&org=NSF (accessed June 2, 2010).

Appendix A. Michigan State University Digital Curation Planning Project Baseline Data Questionnaire

Welcome

Welcome to the Michigan State University Digital Preservation Planning Baseline Data Questionnaire—the first step towards participating in a university-wide initiative that will help you preserve and maintain the accessibility of your unit's data.

1. What is the name of your MSU unit or department?

2. What is your title?

Digital Content

3. What types of digital content does your unit produce? Please check all that apply.
   - Word Processed Documents
   - Imaging—Paper Documents
   - Imaging—Photos
   - Imaging—Non-Photos (e.g., maps, drawings)
   - Digital Photos
   - Digital Graphical Images (e.g., maps, drawings)
   - Audio
   - Video
   - Spreadsheets
   - Databases
   - Presentations
   - Web Pages
   - CAD Drawings
   - Data Sets
   - Other

4. Of the digital content types checked in the previous question, which type(s) make up the largest proportion of the total digital content produced at your unit? Please indicate approximate percentage(s) of total proportion of digital content.

5. Approximately how much digital content does your unit maintain?
(multiple choice, one answer)
   - < 1 TB
   - 1–5 TB
   - 5–10 TB
   - > 10 TB

6. How is your digital content stored? Please check all that apply.
   Hard Drive
   Removable Magnetic Media (e.g., floppy discs, Zip discs)
   Optical Media (CD/DVD)
   Digital Tape
   Solid State (e.g., flash drive)
   Other

File Formats

7. What file formats are created and/or maintained by your unit? Please check all that apply.
   MS Word
   Text
   PDF
   HTML
   TIFF
   JPEG
   WAV
   MS PowerPoint
   MS Excel
   MS Access
   MS Publishes
   Other (please specify)

8. Of the file formats checked in the previous question, which make up the largest proportion of files produced at your unit? Please indicate approximate percentage(s) of total proportion of files.

Technological Infrastructure

9. What is your unit's current storage capacity?

10. Does your unit plan to expand this capacity in the next year?
   Yes
   No

11. If so, approximately how much capacity will be added?

12. Does your unit use any content management or other specialized software systems to manage digital files? (e.g., SharePoint, Luna, Extensis Portfolio, etc.)
   Yes
   No

13. If so, which digital asset management system(s) are used?

14. Does your unit maintain a digital repository?
   Yes
   No

15. If so, what digital repository software is being used? (e.g., DSpace, Fedora, ContentDM)

Confidentiality Issues

16. Is any of your digital content of a confidential or sensitive nature?
    Yes
    No

17. If so, what is the proportion of confidential to non-confidential content?

Contact Information

20. Please provide the following contact information. The MSU Digital Preservation Planning team may contact you shortly to schedule a more in-depth interview.
    Name:
    Email Address:
    Phone Number:

Thank you!

Thank you for participating in this questionnaire. If you have any questions about the MSU Digital Preservation Planning initiative, please contact Lisa Schmidt, digital preservation analyst, at lisa.schmidt@matrix.msu.edu.

Appendix B. Michigan State University Digital Curation Planning Project Unit Interview "Tickler" Questions

Describe the mission of your unit.

Describe your digital content.

How does the digital content relate to the mission of your unit?

What content must be preserved?
    Of ongoing use to unit and/or partners.
    "Archival" in the local sense, documenting the activities of the unit.

Is any of the content archival in the sense that it documents the history of the university and should be in the custody of the Archives?

File formats
    Describe
    Different preservation and access formats?

How stored?

Do they have storage issues?

Discuss CMS and/or DR.

What are they using?

What are they doing with it?

What digital content are they storing in it?

Who uses it?

Why did they choose that solution?

How is it working for them?

Does the system provide preservation functionality, such as checksum calculations?

Are preservation masters stored in the CMS/DR?
    If not, where are they stored?

Are they happy with it, or are they looking at implementing another solution?

Describe workflows
    Ingest
    Archival storage/preservation processes
    Access

Metadata
    Information stored with or related to content
    Any particular metadata schema?

File naming conventions
    Consistent?
    Describe

Appendix C. Michigan State University Digital Curation Planning Project One-on-One Interview Results

Unit and Scope	Content Management System, Digital Repository, and Other Software	Metadata
Confucius Institute at MSU	Alfresco content management system, to manage project workflow (www.alfresco.com)	
Department of Art & Art History	Open-source, web-based Madison Digital Image Database system to manage, share, and organize digital images (http://mdid.org/overview.htm)	Image Resource Information System (IRIS) data standard, uses Visual Resources Association (VRA) Core, Cataloging Cultural Objects (CCO)
Department of Theatre	DOT:Media, a digital repository created using ResourceSpace digital asset management software (www.resourcespace.org);	Based on Dublin Core
MATRIX	In-house developed open-source KORA digital repository software (www2.matrix.msu.edu/research/technology/kora)	Based on Dublin Core
MSU Extension/Agriculture and Natural Resources Technology Services	Knowledge Repository (www.msue.msu.edu/portal), digital repository system custom-developed by Intrafinity (www.intrafinity.com)	Based on Dublin Core
National Superconducting Cyclotron Laboratory (NSCL)	In-house developed, open-source —NSCL Data Acquisition System nuclear physics data acquisition software (http://sourceforge.net/projects/nscldaq); —NSCL SpecTcL Histogramming System, an open-source C++=based analysis package for nuclear physics data (http://sourceforge.net/projects/nsclspectcl)	
Physical Plant Division	—Oracle-based FAMIS enterprise facility management software suite to manage operations (http://solutions.oracle.com/solutions/famis/famis) —InnoCielo Meridian Enterprise software for document management (www.cyco.com/products/ice/) —Skire project management software to track vendor activity (www.skire.com/) —Munsys spatial data management softrare to map underground utilities (www.munsys.com/index.htm) —InStep eDNA Reat-Time Historian to measure utility usage (www.instepsoftware.com/edna_overview.asp)	Local controlled vocabulary for project transactions
Turfgrass Information Center of the MSU Libraries	Cuadra STAR content management system (www.cuadra.com/products/products.html)	Customized indexing terms
University Relations	—Extensis Portfolio media management system for indexing photos and NetPublish Portfolio for public access (www.extensis.com/en/home.jsp) —Zenfolio online delivery system (www.zenfolio.com/)	