5. Supporting Digital Scholarship: Bibliographic Control, Library Cooperatives and Open Access Repositories – Karen Calhoun

Abstract

Research libraries have entered an era of discontinuous change—a time when the cumulated assets of the past do not guarantee future success. Bibliographic control, cooperative cataloguing systems and library catalogues have been key assets in the research library service framework for supporting scholarship. This chapter examines these assets in the context of changing library collections, new metadata sources and methods, open access repositories, digital scholarship and the purposes of research libraries. Advocating a fundamental rethinking of the research library service framework, the chapter concludes with a call for research libraries to collectively consider new approaches that could strengthen their roles as essential contributors to emergent, network-level scholarly research infrastructures.

5.1 Changing collections, the "control zone" and resource discovery

Scholarly collections have changed. Now dominated by licensed online content, academic library collections are becoming more universally available and less institutionally focused. Other types of collections—digital libraries and open access repositories—are gaining in visibility and importance. With the transformation of the scholarly information landscape wrought by the web, it is no longer possible for one individual library to own, license or point to all the information objects of value to the academy. The roles and value of locally housed, largely paper-based collections are being reconsidered (for some recent examples

see Lewis 2013; Malpas 2011; Henry et al. 2011; University Leadership Council 2011; Payne 2007).

5.1.1 The state of the scholarly research collection

The demand for online content has seemed insatiable. In response, starting in the mid-1990s, research libraries began expending at first modest and now large shares of their budgets on commercial online content controlled by publishers. Many foresaw the large impact that the shift to online access would have on library collections and collection-centred services like reference and cataloguing. Ross Atkinson, an exceptional librarian who helped to define modern research library collection development, wrote a seminal paper on the roles of library collections up to the mid-1990s (1996). As he observed more and more scholarly publications moving from paper to online, Atkinson sought to define research library collections in terms of their fundamental purpose "to reduce the time needed by individual client-users to gain access to that information they need to accomplish their personal or institutional work objectives."

In the paper environment, research libraries "privileged" certain information objects by acquiring and transferring them to be physically available to the library's user community—a function "disintermediated" by the displacement of paper with highly-priced commercial online content. Atkinson was an early advocate for research libraries' reclaiming their roles supporting essential components of scholarly research *infrastructure* (the content, technology, tools, services, systems, organizations and facilities on which the continuance and progress of scholarship depend). Since Atkinson published his article, the research library community has vigorously sought ways to carry forward its roles supporting

scholarship through advocacy, open access publishing and the establishment of open access repositories.

While Atkinson's 1996 proposal was more radical than what has since unfolded, his intent was widely shared in the research library community. He proposed reclaiming libraries' central role in scholarly communications through the establishment of a *control zone* (a global digital library, in fact) that would carry forward the purpose of privileging, drawing attention and adding value to scholarly content. He regarded the selection and preservation of the scholarly subset of the universe of information as a social and ethical imperative of librarianship. It is in this context that he argued:

> It is time—past time—for the academic library community to begin work on the creation and management of a single, virtual, distributed, international digital library, a library that has (conceptual, virtual) boundaries, that defines its services operationally on the basis of the opposition between what is inside and outside those boundaries, and that bases that service on the traditional social ethic that has motivated all library operations in modern times. The academic community must consider, in other words, the creation of a control zone. (Atkinson 1996, 254)

He recognized the practical limits of realizing a single, universal control zone and proposed regional and other instances, linked interoperably to a larger whole.

Ten years later, Atkinson (2006) convened an important conference of collection development leaders to consider the future of research library collections. By that time,

while the disintermediation of library services that Atkinson foresaw in 1996 was coming to pass, pieces of the control zone had begun to emerge in unanticipated ways.

5.1.2 Pieces of the control zone on the web

Research libraries have not developed the control zone that Atkinson envisioned, but pieces of it (that is, of a new scholarly research infrastructure) have nevertheless begun to emerge as a result of the development of the web and the tumultuous, technology-driven changes in the practices and tools of scholarship over the past twenty years. Efforts to rethink scholarly communications and advance toward "the system that scholars deserve" (Van de Sompel et al. 2004) are having an impact. Three lines of development—some scholars' establishment of subject-based repositories; research libraries' persistent efforts to establish institutional repositories; and parallel developments resulting from the transformative impact of search engines—have combined to produce a nascent, "network-based system that ... provides interoperability across participating [scholarly communications] nodes."

These developments occurred against a backdrop of changing user behaviours with respect to collections. By the middle of the first decade of the new millennium, students had expressed clear preferences for starting their research with search engines (see De Rosa et al. 2005) and faculty perceived themselves as less dependent on the library (Schonfeld and Guthrie 2007). By the beginning of the second decade of the millennium these trends had become even stronger (De Rosa et al. 2011; Schonfeld and Housewright 2010). For example, across disciplines, faculty members now prefer to begin research with network-level

services (common search engines or discipline-specific online resources) instead of using the library catalogue or visiting the library. Faculty and researchers expect that scholarly content will be online, and it is: the Cox and Cox 2008 study reported that over 95% of science, technology and medicine and 85% of arts, humanities and social science journal titles are now online.

In other developments, search engines have improved their abilities to index the deep web (Zillman 2012), and scholarly content providers are increasingly willing to allow search engines to crawl their metadata or content so that it is easily discoverable there. Also part of the fabric of this transformation is rethinking the role of little-used print collections (for usage statistics see Anderson 2011). In addition to developing storage facilities, research libraries are making their print collections work harder through successful innovations like Borrow Direct, a user-initiated book borrowing programme that essentially makes multiple libraries' print collections available to a larger audience (Nitecki, Jones, and Barnett 2009). The Hathi Trust (hathitrust.org/about) gathers together partner research libraries' collections in digital form for preservation and effective access.

5.2.2.1 Scholarly journals and articles

In practice, Google and Google Scholar are identifying a network-level, scholarly control zone of licenced and open access scholarly articles. Google Scholar has been in place since 2004. The CIBER study (2009) and other research findings (Hampton-Reeves et al. 2009) indicate that both faculty and students now discover online scholarly content more often through Google, Google Scholar and Google Books. Now that the content of scholarly

aggregations (like ScienceDirect and the content of open access repositories) is crawled and indexed by Google, a huge amount of traffic to online scholarly content comes from Google. Google Scholar includes open access versions of articles in search results when possible. Norris, Oppenheim and Rowland (2008) evaluated four search tools' utility for locating open access articles and found that Google and Google Scholar did the best job of locating them.

Metadata for commercial e-journals generally is originally sourced from publishers and passed along in a supply chain to a number of stakeholders including national libraries and registration agencies, subscription agents and vendors. The work of CONSER, the Cooperative Online Serials Program (www.loc.gov/acq/conser/index.html) remains relevant in the metadata supply chain for online journals; however most libraries are more consumers than creators of the descriptive metadata for e-journals.

Metadata for journal articles has more sources. It is generally provided by publishers, abstracting and indexing services, aggregators of online articles, authors and those acting on behalf of authors (such as e-print repository managers or staff). Producing metadata for journal *articles* has never or rarely been the domain of the library cataloguing community. Following the launch of Google Scholar in 2004, more and more article content providers agreed to expose their metadata for harvesting by search engines and for inclusion in the central indexes of library discovery services.

5.2.2.2 New library discovery services

By tradition, discovery of articles has been the domain of abstracting and indexing agencies and bibliographers, not cataloguers. Article-level metadata has typically been absent from

the catalogue. With the massive changes that led to the direct online delivery of articles, and the high user demand for this content, it became critical for the catalogue to somehow integrate and surface article-level content. Starting around 2005, new discovery services began to replace traditional library online public access catalogues (OPACs), because the traditional catalogue could not scale to this purpose (for example, as of this writing the author's ARL library's discovery service indexes over 71 million journal articles).

The new discovery services (Summon, EBSCO Discovery Services, WorldCat Local, Primo, Encore, etc.) provide a common interface to a centralized index of pre-harvested, pre-indexed metadata or content from heterogeneous sources, thus enabling a library to present a much larger view of scholarly content to its users. In additional to indexing e-content (journals, books and articles) to which the library has access, typically the discovery service also indexes the library catalogue, institutional repository and digital collections. See Chapter 4 for more information.

5.2.2.3 Open access repositories

Open access repositories grew out of the desire to transform scholarly communications and reduce dependence on commercial publishing. They exemplify what Atkinson had in mind when he envisioned "reclaiming" the control zone from publishers. On balance, they have been successful: a directory of open access repositories, OpenDOAR (opendoar.org), lists nearly 2,200 repositories, a figure that has been steadily rising each year.

Repositories, both subject- and institutionally-based, are gaining in visibility and impact. Several of the subject-based repositories have succeeded in transforming scholarly

communications and fostering worldwide collaboration in the disciplines they support. Institutional repositories have been challenged with low deposit rates, but they are gradually becoming more successful at attracting submissions. Deposit mandates (governmental or institutional requirements that researchers make their papers available in open access repositories) have begun to stimulate growth in the number of open access papers available (see for example publicaccess.nih.gov). The result is that the number of articles published in open access journals or available from open access repositories (both subject-based and institutional) or authors' web pages represents an increasing proportion of annual scholarly output. Björk and colleagues conducted a number of analyses (Björk, Roos, and Lauri 2009; Björk et al. 2010; Laakso et al. 2011) and estimated that a little over 20% of all the articles published in 2008 were openly available a year later. The results of the study that the team published in 2011 suggested that the number of open access journals has grown at an annual rate of 18% since 2000, and the number of open access articles has grown 30% a year.

Open access repositories have become increasingly important for improving the discoverability and accessibility of not only articles, but also reports, theses and dissertations, conference and working papers, teaching materials and presentations. Universities have had some incentive to invest in repositories as ways of better organizing and disclosing the scholarly output they support.

To provide a sense of the scope and scale of attention received by selected repositories compared to other sites of interest to scholarly information seekers, figure 5.1 compares US traffic in October 2011 across a number of sites including:

- A highly-ranked institutional repository (dspace.mit.edu)
- Three of the top subject-based repositories
- Several library-related sites (Yale, the catalogue of the Library of Congress, the University of Michigan, New York Public Library and OCLC's public interface to WorldCat)
- A leading commercial site for science, technology and medicine (STM) research (ScienceDirect)
- Google Scholar

Numbers of unique monthly visitors to the repositories are shaded in grey.

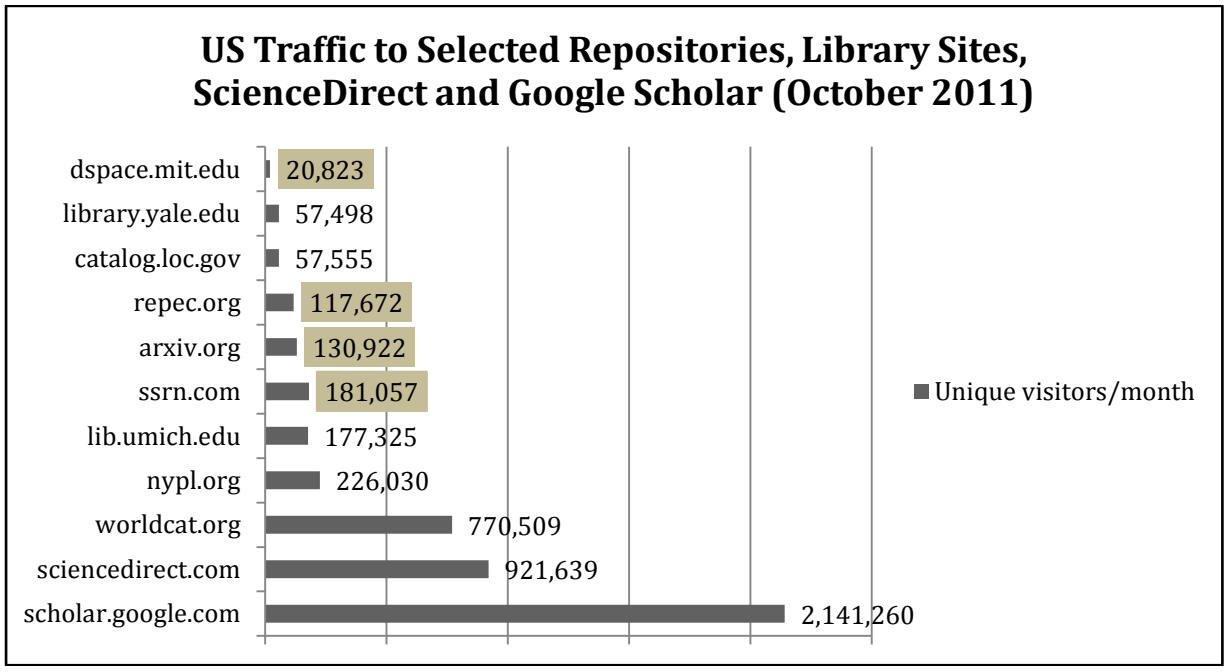

Figure 5.1. US Traffic to selected repositories, library sites, ScienceDirect and Google Scholar (October 2011)

Source of data: siteanalytics.compete.com

The metadata for repositories generally comes from authors or repository managers on behalf of authors. Sometimes metadata librarians or cataloguers participate in this process. Authors and subject experts have proved to be good sources of metadata (see for example

Wilson 2007), especially when the author or contributor interface provides helpful guidance. A variety of investigations (for example see Duranceau and Rodgers 2010) are looking into how metadata might be more easily generated and transferred to multiple repositories at the point the content is deposited.

5.2.2.4 Digital libraries

An analysis of digital libraries and digital library aggregations suggests the following categories of digital objects (figure 5.2).

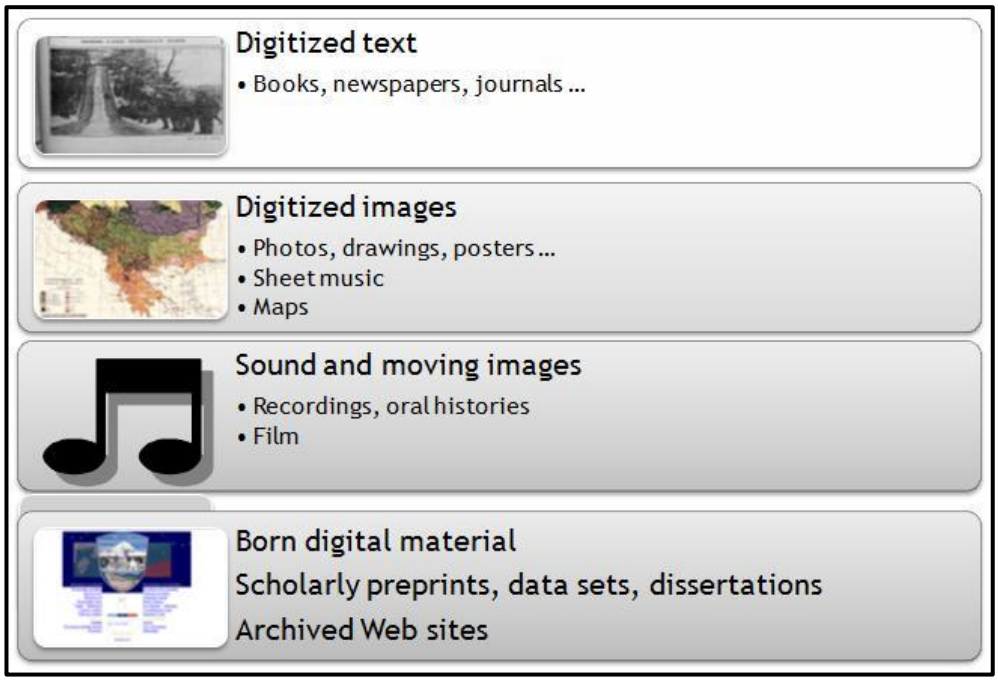

Figure 5.2. Types of digital objects

There are indications that larger digital library collections are attracting a good deal of attention. Figure 5.3 is a chart using 2012 data from Alexa.com showing the percentage of visitors who go to different parts of a site. The chart shows the top three visitor destinations for the national libraries of France, New Zealand and Australia. Gallica, PapersPast and Trove are digital libraries.

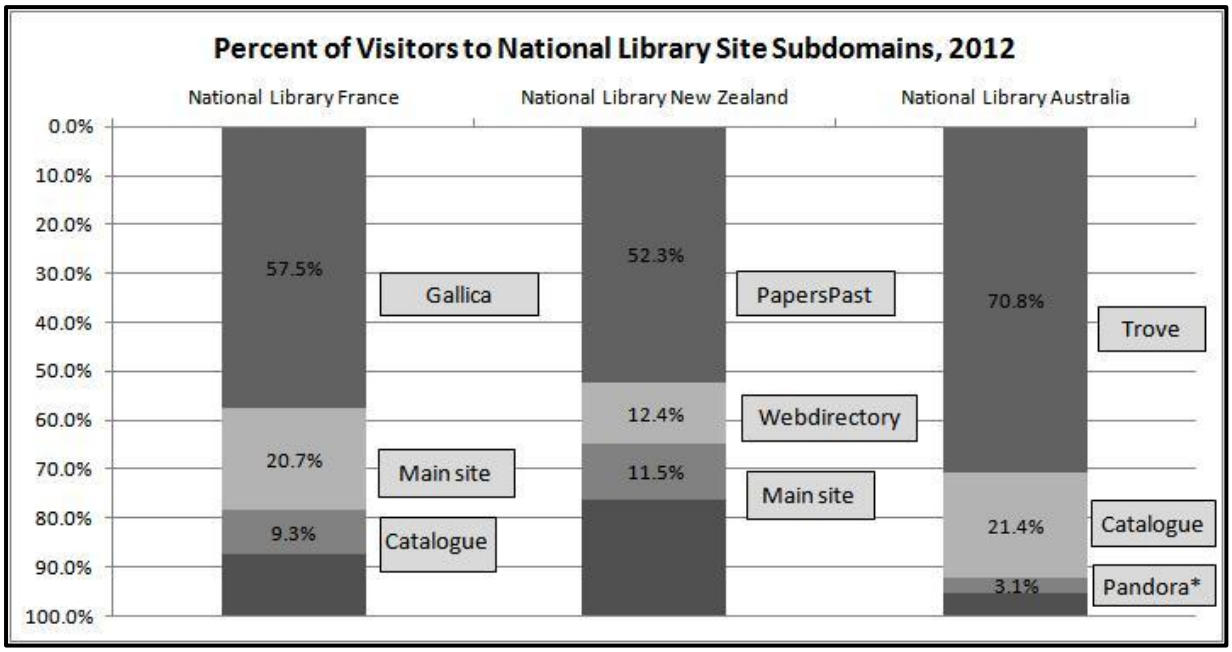

Figure 5.3. Percent of visitors to digital library subdomains on national library sites, 2012

Source of data: alexa.com

***Pandora is Australia's web archive**

The metadata for digital library collections comes from many sources and different metadata communities, which deploy a variety of content rules, frameworks, schemes, reference models, element sets and encoding rules. NISO's *Understanding Metadata* (Guenther and Radebaugh 2004) provides helpful guidance for anticipating the scope, complexity, and interoperability requirements for representing large-scale digital libraries of cultural heritage materials. While many digital library sites are crawled by Google and are thus discoverable through search engines, their easy discovery alongside other types of content of interest to scholars and students remains problematic. The new library discovery services (Summon, Primo and the like) can sometimes index the metadata from digital library collections so that these collections are discoverable from a common library interface.

5.2.2.5 Books

The transformation of discovery and access to the scholarly journal literature progressed fairly rapidly from the mid-1990s until now. The transformation of discovery and access to books is progressing more slowly, but developments over the last decade have continued to alter the landscape of book collecting and metadata for many types of books, for example:

- The development commencing in 1999 of the ONIX (Online Information Exchange) standard for representing and communicating book industry product information in electronic form. The importance of ONIX has grown among publishers, booksellers and other participants in the book industry supply chain. Luther's white paper (2009) explores the possibilities for moving to an environment where metadata is exchanged more easily throughout the supply chain. In the process of her investigation she prepared an excellent Book Metadata Exchange Map which is well worth the reader's time to study. The map illustrates how publishers, metadata aggregators, wholesalers, booksellers (e.g., Amazon), Google, buyers and readers, and national/local libraries participate in the books metadata supply chain. Libraries are one of many participating types of organizations.

- The success of Amazon, founded in 1994, which has raised reader expectations for books metadata creation and management well above what libraries can accomplish with library-based methods (for a comparison of library and Amazon metadata see Calhoun 2011, slide 23).

- The availability of machine-to-machine web services (e.g., from Amazon, Google Books, Ingram, etc.) that enable web developers to easily capture metadata for books for re-use in their own sites.

- The latest figures suggest that a total of 22 million books may have been digitized in various projects, the majority by the Google Books project (estimated September 2012 from figures at sustainablecollections.com). Hathi Trust provides APIs and other methods that web developers can use to retrieve metadata for a subset of these titles (hathitrust.org/data).

- The gradual acceptance of commercially produced e-books. With the rising popularity of e-book readers and tablets, the pace of e-book acceptance has quickened in the last two years. In 2012, Amazon e-book sales exceeded sales of both paperback and hardcover books (Malik 2012). The metadata for these e-books is often produced outside the library cataloguing community and adapted by vendors for use by libraries.

Library cataloguing's role has shifted from a central position to one of many methods supporting resource discovery and delivery on the network. It is important to point out some exceptions to the trends in how mainstream published books and e-books are selected, packaged, and delivered to libraries: metadata creation and dissemination for non-English text, music, video, maps, special collections, images, archives, and new formats will likely continue to be resistant to these trends.

5.2.3 Special collections and archives

As more scholarly content moves online and academic libraries license the same or similar e-content packages, individual libraries' online collections become less distinctive. There is also considerable overlap in many legacy print collections. Special collections and archives are what remain most distinctive about library collections. These types of collections are far from being part of the nascent research infrastructure that is emerging on the web. The results presented in a previous section of this chapter suggest that, if special collections of cultural heritage materials were more discoverable online, they would be heavily used.

Unfortunately, improving discoverability and access to special collections and archives is difficult, because many of them are hidden (either not catalogued at all or not represented in online catalogues). A UK study of hidden collections (Loughborough University. Library & Information Statistics Unit and Research Information Network 2007) found that half of UK research library collections were not represented in online catalogues at the time of the study. Given today's information seeking preferences, this means these materials may as well not exist. The study's conclusions urged a UK-wide strategy and programme to uncover these hidden collections.

Based on the number of hidden collections documented by a 1998 survey by the Association of Research Libraries (ARL), an ARL task force worked from 2001 to 2006 to advance a seven-point action plan to promote and surface hidden special collections in US libraries (Association of Research Libraries 2008). Subsequently, in 2008 the US Council on Library and Information Resources (CLIR) began to administer a national effort with the support of

the Mellon Foundation, awarding grants to support efficient description and processing of large volumes of hidden special collections material of high value to scholars (clir.org/hiddencollections). OCLC Research conducted a study in 2009 to evaluate progress since the 1998 ARL survey (Dooley and Luce 2010). They found some progress, but not enough. Many still lack any online representation.

5.3 Metadata, cataloguing, bibliographic control

5.3.1 Change

Metadata has changed as collections have changed. It remains important, but it comes in many forms and from many sources. The centrality of bibliographic control has been disrupted. It has become just one piece of a large puzzle that must be assembled to provide for discovery and access to changing collections. Discovery and access for online and print books and journals and scholarly articles are supported through metadata supply chains in which libraries are one of many participants. Metadata for providing aggregation, discovery, access and management of e-content and the contents of open access repositories and digital library collections increasingly come from outside traditional library cataloguing environments and cooperative cataloguing systems. As previously noted, there are some exceptions to these trends: (1) currently received non-English texts, music, video, special collections, images, archives and new formats; and (2) hidden collections.

5.3.2 Metadata management

The traditional manual methods of bibliographic control employed in libraries have not easily scaled as scholarly content moved online and user demand for electronic resources

grew. A transition to a metadata management approach began in the 1990s, featuring metadata reuse, automated methods, and distributed participation in metadata creation and enrichment, including authors and end users. The new library discovery services are another indication of the transition from bibliographic control to metadata management: they create a common interface through a central index to heterogeneous sources. The most recent example of a metadata management approach is PDA (patron-driven acquisitions), a vendor-assisted workflow that obviates the need for local cataloguing services for these books. Metadata management approaches are variously supported in current cooperative cataloguing systems.

Table 5.1 compares bibliographic control and metadata management practices. In large research libraries, the shift to metadata management (in line with the changes in the right column of the table) has generally occurred in parallel with, rather than as a replacement for, the traditional workflows of bibliographic control. In her extensive literature review of the cataloguing literature of 2009 and 2010, Gardner (2012, 68) poses the question "To what extent has the cataloging community embraced non-MARC metadata creation and interoperability?" and answers "not much."

Wolven (2009), pondering the waning role of traditional cataloguing as the library moves to the web, nevertheless expresses optimism that the library community will "reach a new consensus on best practices, less grounded in 20th-century publishing patterns." As the cataloguing community's debate about bibliographic control rages on, downward pressure on library budgets, shrinking cataloguing departments, the search for efficiencies, and

opportunities for new ways of doing things have continued to propel the shift from traditional cataloguing toward metadata management.

Table 5.1. Toward distributed systems for metadata management

Bibliographic control	Distributed metadata management
• For finding and managing library collections (mostly print)	• For finding and managing many types of materials, for many user communities
• Catalogue records (well-understood rules and encoding conventions)	• Many types of records, many sources, disparate treatment
• Shared cooperative cataloguing systems	• Loosely coupled metadata management, reuse and exchange services among multiple repositories
• Usually handcrafted, one record at a time	• Multiple batch creation and metadata extract, conversion, mapping, ingest and transfer services
• Record creation and editing generally a solitary activity undertaken by library-trained professionals	• Distributed metadata creation and manipulation; dynamic records; digital library metadata - a fundamentally collaborative activity involving various specialists

5.3.3 Responses to disruption

Since the new millennium, the disruptive new conditions brought by the internet and web have eroded prior certainties and produced an entire body of reports and articles around the future of catalogues and cataloguing. On behalf of the Library of Congress (LC), this author produced one of these reports (Calhoun 2006). A number of extensive investigations have produced recommendations (US and UK examples include Library of Congress. Working Group on the Future of Bibliographic Control 2008; Reynolds et al. 2009; Chad 2009). Practicing cataloguers, technical services managers, and other library professionals have expressed a range of views around implementing the recommendations of these studies, from strong support and calls for action (for example Hruska 2009; Hruska 2011) to

fervent opposition (for example Bade 2009). The community discussion so far has not created consensus.

The long development and implementation path of a new cataloguing code, Resource Description and Access (RDA), has not clarified the way forward. At least in the US, RDA has been a topic of debate. The recommendations of the US RDA Test Coordinating Committee stated "the test revealed there is little discernible immediate benefit in implementing RDA alone. The adoption of RDA will not result in significant cost savings in metadata creation. There will be inevitable and significant costs in training." The committee reported the results of a survey of US cataloguers that indicated mixed support for implementation of RDA (Cole et al. 2011, 4–5). For a variety of reasons, the committee did nevertheless recommend that the three US national libraries move ahead with RDA implementation, provided some key issues could be addressed and that implementation begin no sooner than early 2013.

One of the Test Committee's pre-RDA implementation requirements was "demonstrate credible progress towards a replacement for MARC. " The committee urged the development of a more flexible cataloguing standard that would address the needs of a changing digital environment. In 2011 LC announced it would work with Library and Archives Canada and the British Library on an initiative whose focus would be "to determine a transition path for the MARC 21 exchange format in order to reap the benefits of newer technology while preserving a robust data exchange that has supported resource sharing and cataloguing cost savings in recent decades" (Library of Congress 2012a).

LC later announced a Bibliographic Framework Initiative to "translate the MARC 21 format to a Linked Data (LD) model" (Library of Congress 2012b). The LC initiative follows an initiative by the British Library to release the British National Bibliography as linked data (Wallis 2011), with the intent to provide opportunities for experimentation among those wanting to interact with bibliographic data in ways that are aligned with web-based practices. While the intentions of these new initiatives are laudable and hold promise for improving the utility and reusability of legacy collections metadata, their focus appears to be limited to how bibliographic and authority data will be modeled, structured and encoded for distribution and consumption on the web. More change is needed than replacing MARC with a linked data model. The Framework Initiative as well as the work of the W3C Library Linked Data Incubator Group (www.w3.org/2005/Incubator/lld) (See Chapter 8 for more information) appear to be centred primarily on standards, metadata, technical issues and advocacy for a linked data approach, while the burning discovery and access issues for research libraries are social (How do scholars and students find the information they need? What roles do and should libraries and library systems play in how they accomplish their work?), economic (What are the costs, what are the benefits of various choices going forward? Who pays?) and organizational (How should research libraries be structured to provide for discovery and delivery of content? What partners do and should they have?).

5.4 Cooperative cataloguing

5.4.1 Foundations

The cooperative model of bibliographic control has been tremendously successful at supporting the development of local library collections and access to these collections. Early automated cooperative cataloguing systems carried forward 1960s and 1970s requirements for collection building and access. This was a time when collections were on site and catalogues were (necessarily) separate from but physically available in the same space as the collections. There was no doubt about what the "collection" was because it was in the building (or buildings), and there was little doubt about what the local catalogue needed to describe.

5.4.2 Cooperative cataloguing today

Cooperative cataloguing databases reflect what library collections once were, not what they have become. For example, the content of the bibliographic database that supports OCLC cooperative cataloguing services represents mainly print-based collections. Most (80 to 85%) of the records in OCLC's bibliographic database describe books (figure 5.4), and while the proportion of e-book records has grown in recent years, over 90% of the books records still represent printed books (OCLC 2012, 16). Schuitema (2010), in her review of the history of cooperative cataloguing, concludes that the shared catalogues that libraries have been producing "while still scalable in terms of providing access to print materials if significant changes were made, will no longer meet the discovery needs of our clientele seeking information in today's expanding digital environment." A helpful discussion of the OCLC bibliographic database's representation of digital content is available (Lavoie, Connaway,

and O'Neill 2007); however the number of records describing digital content has until recently grown more slowly than searchers' appetites for digital content.

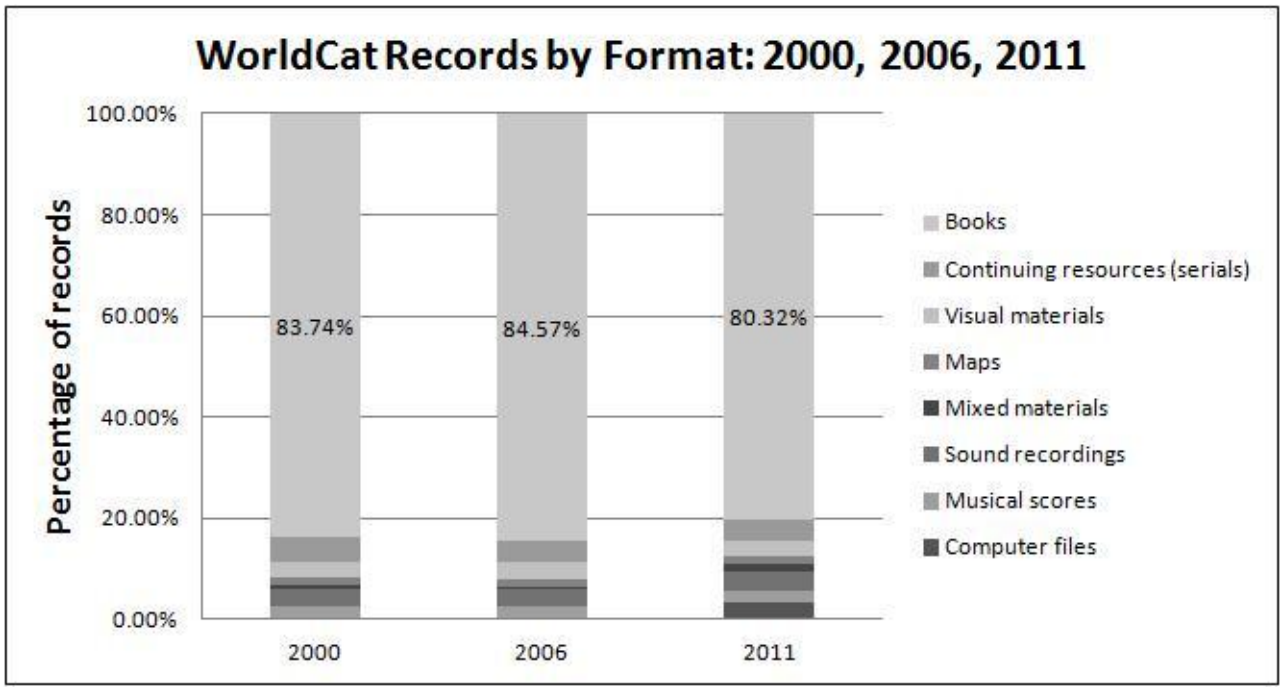

Figure 5.4. OCLC WorldCat by type of material described, 2000, 2006, 2011

Source: OCLC annual reports

(oclc.org/news/publications/annualreports)

5.4.3 Stores of MARC records

Cooperative cataloguing systems serve principally as sources of MARC records that are copied and reused in local library systems and institution-level catalogues. Several writers (for example Chad 2009; Eden 2011; Sellberg 2010) have questioned the necessity of continuing with this model, which exists to facilitate the duplication of cataloguing data. As scholarship moves to the network level and research library collections change, and as metadata becomes available from numerous sources, the central position of local catalogues produced using traditional cataloguing practices is diminished. Many of the new metadata management approaches described earlier, while making local processing more

efficient and less costly, merely perpetuate the current model, which is based on duplicating MARC records in local catalogues.

When MARC and cooperative cataloguing databases of MARC records first emerged, it was necessary to copy the machine-readable records onto some type of physical media—usually cards—because at that time libraries had no online systems in which to store them. Multiple local catalogues representing multiple local collections made sense in the print environment, and the duplicative costs of building and maintaining them were justifiable.

In the mid-1980s, when libraries began to move away from card catalogues toward local online systems, the same practices were carried forward: upon acquiring a new item for the collection, search for a cataloguing record, then copy it for the local system (rather than order a set of catalogue cards). Once a cataloguing record had been obtained, the rest of the work to support discovery, access, and collection management has tended to be done at the institutional level. This set an important precedent and operational model which persisted even as collections moved online in the 1990s; in keeping with the local catalogue model, libraries generally provide for e-content discovery and management locally and redundantly rather than at the network level.

A recent survey of North American MARC record providers, commissioned by LC, provides a fairly complete picture of who produces and distributes MARC cataloguing records in the US and Canada (Fischer and Lugg 2009) for re-use in local systems. By far the largest North American suppliers are OCLC and LC. The results of Fischer and Lugg's survey also reveals the extent to which vendors have entered the MARC record supply business, notably as

sources of MARC records for books but also new types of library materials (e.g., e-books). Ken Chad's study (2009) reveals the flow of bibliographic records and the role of vendors and shared cataloguing aggregators in the UK.

5.4.4 Redundant catalogues

Ken Chad examines the distinction between redundant *cataloguing* (re-editing records to suit local practices) and redundant *catalogues.* Pointing out the duplication of effort required to produce and maintain multiple individual library catalogues, he enumerates the benefits of moving from the UK higher education sector's 160 standalone catalogues to a single shared catalogue at the network level for all of these libraries. Eden (2011) has urged the cataloguing community to turn decisively from its past focus on local operations toward working at the network level.

An option is to move from multiple standalone catalogues to larger shared, cooperative frameworks at the network level that would register many libraries' holdings and be able to feed this information to multiple locations on the web. Such frameworks might operate at a global level or as participating nodes at local, regional or national levels. Roxanne Sellberg (2010), in an article on the future of cooperative cataloguing, writes:

> The identity of individual libraries has traditionally been based on locally defined, but highly duplicative, collections. The need to build equally duplicative catalogs has provided incentive for cooperative cataloging. In the future, the collections of libraries will be not so much duplicated as shared, and separate libraries may not

have separate public catalogs. Is there a place for cooperative cataloging in such a future?

Sellberg observes that collections are now more web-based and advocates that the library community work together to build a worldwide shared catalogue on the network. This worldwide catalogue would be fundamentally different than today's cooperative cataloguing databases. Her recommendations for creatively reshaping cooperative cataloguing systems include:

- Reorganizing such systems around metadata that individual libraries can point to (or from) instead of copy
- Providing better functionality to enable the contribution of metadata for rare and unique library holdings
- Offering additional support for contributing metadata for born-digital information resources not well covered by other players in the scholarly information space

5.5 Rethinking

5.5.1 Bibliographic control

Scholarship has moved online, and the universe of information objects of potential interest to scholars and students can never again be captured as part of a single research library's collection. In this fundamentally changed world of scholarship, the role of bibliographic control has been diminished and marginalized. There is less need and place for traditional bibliographic control as a set of methods for providing for discovery, access and management of the content of mainstream books and serials. For other types of digital

objects, there are other sources of metadata (or other techniques based on crawling and/or indexing) that are better suited for and less costly to use in a web environment.

Even if other sources of metadata and techniques were not available, the manual and semi-automated methods of bibliographic control have not and will not scale to the size of the new universe of scholarly information objects. It is unreasonable to think that even enhanced methods of bibliographic control could support the discovery, delivery and management of this wider array of objects. As for legacy bibliographic data, experimentation with repackaging and distributing it as linked data for consumption by web services is of interest and worth pursuing; however such experimentation is unlikely to address the fundamental issue that going forward, the techniques of bibliographic control have diminished utility in a transformed world of research and learning.

The library community already takes advantage of other metadata supply chains for journal articles and other types of scholarly information objects. This chapter suggests that the library community consider the feasibility of turning away from the traditional methods of bibliographic control for mainstream books and journals in favour of methods used in the book industry and scholarly publishing.

Bibliographic control would continue to have a role: creating metadata for unique content that is unlikely to be described any other way. As noted earlier, some types of content are resistant to the trends around metadata for mainstream published books and journals (electronic and print): for example non-English text, music, video, maps, special collections, images and archives. In addition, many hidden collections containing unique content remain

to be described using more traditional methods of bibliographic or archival control. It would be tremendously useful if effort now tied up in building and maintaining redundant local catalogues could be freed up and systematically applied to improve the discoverability and management of these types of unique non-digital content.

5.5.2 Registries of library holdings

Arguably, what is most important about the data stored in local catalogues is the holdings information—not the bibliographic data. In information systems consisting of brief metadata disconnected from content (such as bibliographic descriptions), what the information seeker usually wants to know is the location of the content and how to get it. This chapter proposes that the library community consider the feasibility of a set of network-level registries of library holdings information that would serve the functions now provided by local catalogues. This set of registries could allow the library community to free itself from the necessity of redundant catalogues (and redundant discovery layers that rely on the existence of copies of bibliographic records in redundant integrated library systems).

Registries are already widely used on the web. Google Scholar is already close to functioning, in practice, as a "registry" of library holdings for scholarly articles. Further, OCLC and Google already have an agreement to provide a "Find in a library" service from Google Books. This service needs further enhancement to make it more functional, usable and visible, but in effect it is a registry of many libraries' holdings of books. E-journal "A to Z" lists are produced from a data set or knowledge base that functions as a registry of the e-journals to which a particular library has access. Through collaborative action and partnerships, research libraries could feasibly move to registry-based systems. While there

is some risk that Google's interest in Google Scholar and Google Books will not persist, it would be a pity if library-supported registries were not integrated with what Google Scholar and Google Books already do. These registries would point from cloud-based discovery services to what libraries hold (or license, or want to point to), as suggested by Sellberg. Web services-enabled registry services could be the basis of a range of cloud-based library services that would obviate the need for copy cataloguing as well as the local catalogue in its current form. Instead of downloading a copy of a record (or obtaining copies of records for loading into a local system) as libraries do now, they would register their holdings of new titles (or sets of titles) that they want to make accessible to the communities they serve.

5.5.3 Standards

Collections of value to scholarship (and their metadata) are diverse and widely distributed in multiple digital and non-digital stores. Libraries have generally approached the challenge of providing coherent access to diverse, distributed content by agreeing on and implementing library-specific standards and using metadata to achieve integration and enable data exchange. These measures have worked well to the degree that everyone complies with the standards; implements them in the same way; and not only uses metadata, but also uses types of metadata that can be cross walked or converted. These methods have not been widely adopted outside the library community. They have been perceived as more complex than the generally simpler, lighter-weight approaches used by many web developers (and increasingly by open access repository developers, as will be discussed shortly).

5.5.4 Building for the web

It is possible that the time has come for new approaches that are built on and for the web. The new global research infrastructure that is emerging to support digital scholarship operates on the network; it deploys standards, protocols and methods native to the internet and web. Any new research library service frameworks would need to be built the same way. Such frameworks need to be web services-enabled, support collaboration and user community engagement, and provide for easy and open data reuse and sharing. Semantic web approaches may provide part of needed solutions and tools, but all hopes should not be pinned on them, as they are relatively untested. A number of avenues, methods and means should be explored.

5.5.5 Enlarged roles for open access repositories

Open access repositories have become increasingly important for responding to changes in scholarly communications and improving the discoverability and accessibility of scholarly information. In keeping with US and other national policy changes related to publications from grant-funded research, they are expected to become increasingly important components of research infrastructure (Borgman 2007, 243). They have potential for becoming key building blocks in frameworks supporting e-research data (for further discussion of e-research and repositories see for example Michener et al. 2011; for research library roles in e-research and data management see for example Soehner, Steeves, and Ward 2010)

Emerging, next-generation open access repositories are built on and for the web. They manage, reuse, repurpose and share/disseminate heterogeneous content and metadata. Interesting starting points for further discussion might be generated from work being done in the Hydra Project (Green and Awre 2009) and to advance ORE (Object Reuse and Exchange; Lagoze et al. 2008; Witt 2010) and SWORD (Simple Web-service Offering Repository Deposit; Lewis et al. 2012). Both the Hydra Project and ORE feature semantic web methods (data models and encoding). The Australian Research Data Commons (described at ands.org.au) is another potential source of ideas: the Commons enables flows for depositing and disclosing content and metadata for re-use and integration in a variety of locations: institutional and domain-specific portals, national services and search engines (Burton and Treloar 2009).

The contents of repositories are usually crawled by search engines and so are highly visible on the web. Google Scholar indexes and provides links to both licensed and open access versions of articles in search results when possible. The potential and feasibility of replacing the current redundant framework of multiple local catalogues with a network of better integrated open access repositories, tied to a larger cooperative infrastructure of "participating nodes," (as defined by Van de Sompel et al.) should be explored. The research library community and its partners have the opportunity to consider substantially different, collectively supported, network-level services that enhance the positive impact and visibility of scholarly content, using a set of loosely connected participating nodes—open access repositories—as a starting point. A later section and figure 5.5 explore one possibility for how such a framework of services might be tied together.

5.6 A library cooperative commons

The notion of displacing multiple library catalogues with a more centralized system or radically different frameworks has been suggested before (for example Coffman 1999; Bell 2003; El-Sherbini and Wilson 2007). Coffman's proposal for "earth's largest library" in particular set off a storm of controversy (see for example Napier and Smith 2000). Yet it seems important to consider at this point that local and union library catalogues are not ends in themselves; they are means. When collections change significantly, and information seekers preferences change, it is necessary to evaluate the means for managing collections and providing discovery, access and preservation. As for research libraries, it seems important to articulate a new, shared strategy for their roles supporting scholarship and learning, a role that otherwise seems well on the way to disintermediation.

Perhaps conditions have now changed enough for research libraries to begin a new conversation. The global scholarly research infrastructure is emerging; however progress toward aligning research library service frameworks with the new scholarship has been largely incremental and reactive. Nevertheless pieces of a new research library service framework already exist, and it may prove feasible to further develop this framework using registries and other web-based techniques. The purpose would be to enable the drawing of virtual boundaries on the network that define "holdings" (digital and non-digital, licenced and open access) of interest to scholars and students. These are not wholly new ideas. In particular, Dempsey anticipated by years the immense impact of network-level discovery environments on local and shared library catalogues. Among many other thoughts about how to evolve library systems in a network environment, he expressed early interest in

system-wide registries as ways of re-connecting network-level discovery to library fulfillment without placing additional burden on local library development (Dempsey 2006, under “Routing”).

Much attention has recently been focused on the possibilities for cloud-based library systems that will allow libraries to replace their local library systems with web-based applications that are accessed via common web browsers and whose infrastructure is supported in the cloud (See Chapter 4 and Breeding 2012 for more information). Since these offerings are so new, it is difficult to predict how libraries will respond, but it seems clear that there will be a role for cloud-based platforms in renewed frameworks for library operations and library cooperative cataloguing. One question is whether the cloud-based library systems now in development, testing or early adoption phase have been or are being designed with the changing requirements of digital scholarship in mind. It would be a missed opportunity if they merely carry forward current library system functionality (acquisitions, inventory control, collection management, holdings maintenance, etc.) for mainly online and print book and journals plus licenced aggregator packages of licenced articles.

5.6.1 A recombinant, cloud-based service framework

Lavoie, Henry and Dempsey (2006) wrote of the need for a “service framework” or shared view of how library services should be organized in a radically changed information landscape. They called for “reusable, recombinant, and interoperable library services.” Subsequent sections of this chapter lay out some possibilities for two components of a new

service framework for consideration. The first component is a set of recombined, interoperable open access repositories managed by research libraries; these would operate as "participating nodes" in the framework. The second component is a cloud-based set of services and registries—one or more library cooperative commons platforms?—that provide interoperability across participating nodes at local, regional, national and global levels. The intent of both is to create a new, network-level research library service framework that aligns and integrates well with network-level scholarly infrastructure and better supports research library roles in 21st century research, teaching and learning.

Why a new role for open access repositories? Open access repositories already operate on the network and their contents and metadata are crawled and indexed by search engines. Next-generation open access repositories are beginning to emerge. Fundamentally, their purpose is to support scholarship as it is now practiced, disseminated and discovered on the web. They are designed to support researchers and to engage the scholarly communities they serve, so they have both user and machine interfaces to support deposit (including deposit in multiple stores of content and metadata). Research libraries might also use them (some already do) to enter the content and/or metadata for digitized, digital or non-digital special collections—instead of putting the metadata in local catalogues or other systems that can't be crawled, harvested and integrated in other discovery services. Metadata for unique hidden collections could also be ingested or entered manually in repositories for disclosure and registration.

As envisioned, in addition to open access repositories, new library cooperative commons services would be essential components of a cloud-based, virtual research library service framework. The commons services would help to tie the repositories loosely together, ingesting unique metadata or content, registering library holdings or other information, and dispersing and/or exchanging data with other nodes or services in the framework.

Considered as a whole, the service framework's purposes would include promoting the sharing of knowledge; enabling researcher and student engagement, participation, content exchange and collaboration; and reducing the time required for scholars and students to identify and gain access to needed information. The objectives of such a framework would include helping research libraries to progress open access to scholarly outputs; extend their institutional roles supporting the infrastructure of digital scholarship; promote the visibility and use of unique collections; and free up local resources for investment in new priorities.

5.6.2 One possibility among many

Figure 5.5 attempts to provide a functional view of one possibility for a new research library service framework from the perspective of one "participating node" (repository). The figure attempts to depict a hypothetical framework that loosely connects and recombines locally-created content and metadata deposited in next-generation, institutionally-based open access repositories with other aggregations/repositories at various levels (domain-specific, regional, national, international); search engines; and one or more library cooperative commons. These local next-generation repositories would take the place of local catalogues and expand the catalogue's functions. Within this hypothetical framework, unique (new)

content and/or metadata can come from content creators or their agents (one type of agent could be a cataloguer or metadata specialist) or from a variety of feeds from other repositories or aggregations. A current example of this kind of workflow is described by Duranceau and Rodgers (2010), who report on an experiment to enable automated deposit of their university's faculty papers into the university's institutional repository when these papers are first deposited elsewhere.. The locally-managed repository would also support the registration of library holdings in a variety of network-level services, including one or more library cooperative commons.

In this hypothetical framework, the role of a library cooperative commons is to provide cloud-based services supporting:

- Local library collection-centred activities (acquisitions; inventory control; licensing; registration of holdings, e.g., for non-unique content; circulation; preservation; digitization)
- Services that promote the visibility and reuse of digital scholarship (including for example e-research data)
- Services and registries that tie the infrastructure together, provide interoperability, and integrate work done across participating nodes (e.g., other locally-managed repositories)

Registries and identifier services would play a key role in this hypothetical framework, as they would be the means for (1) registering library holdings, thus enabling libraries to move

toward *sharing* metadata instead of *copying* it locally, (2) tracking hidden collections and promoting their processing and (3) providing more efficient means for linking and pointing to content. There could be other types of registries or knowledge bases in addition to registries of library holdings, as discussed by Dempsey (2012).

The idea for this figure came from an illustration describing the infrastructure supporting Research Data Australia, a network-based service that connects data, projects and researchers across organizations and promotes the visibility of Australian e-research data nationally and in search engines (ANDS 2012). Figure 5.5 is not an architectural view and makes no attempt to comment on how content and metadata are actually stored or how the infrastructure works from a technical perspective. It merely sketches one hypothetical possibility from an institution-level processing or workflow point of view.

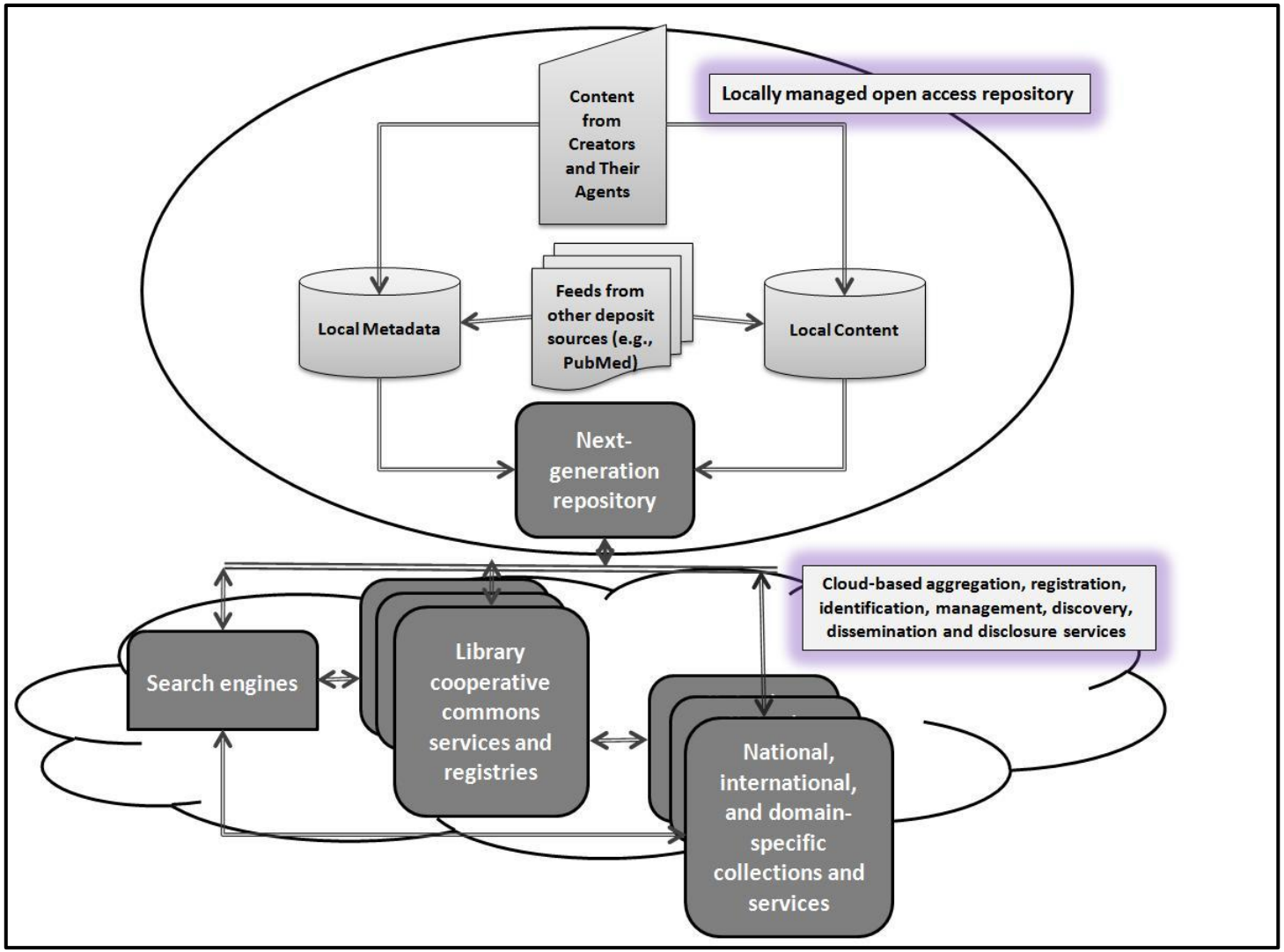

Figure 5.5. A hypothetical service framework for research libraries

5.7 Conclusions

Atkinson's control zone has not and will not emerge as he anticipated it could. Nevertheless pieces of the control zone have come to exist in the form of

- Google Scholar and Google Books;
- An informal network of subject-based and institutional open access repositories; plus
- A number of national and international-level digital libraries of cultural heritage materials

It is not too late for research libraries to lay claim to the scholarly control zone—or at least to participate more strategically and intentionally in the pieces of the network-level scholarly research infrastructure that have already emerged. Should research libraries and their partners choose to engage collectively in these emergent research infrastructures, they could reclaim now fully or partially disintermediated roles supporting research, teaching and learning, for example:

- Promoting the visibility and recognition of valuable collections and scholarship on the web
- Engaging more actively with scholars and students
- Reducing the time required to identify and gain access to needed information

This chapter advocates a fundamental rethinking of the apparatus (service framework) that research libraries now deploy to support scholarship. Today's local library systems and catalogues, regional and union catalogues, and cooperative cataloguing systems collectively constitute key components of a service framework that was useful when print collections stood at the centre of scholarship and traditional library methods for providing discovery

and access were sufficient. To continue to focus solely on improving this long-successful service framework (and on bibliographic control alone) implies a choice to limit the role of research libraries to supporting a subset of the information environment, leading eventually to a diminished role in higher education, scholarly communication and cultural heritage.

Open access repositories, while they have their own challenges, look to the future and have great potential for directly engaging scholars and students and for supporting scholarship as it has become. This chapter calls for an exploration of the possibilities for a new, network-level research library service framework—one that features a substantially larger role for open access repositories. This chapter proposes that "catalogue 2.0" is not a catalogue at all, but a participating node (a repository) in a new library service framework for supporting scholarship and learning.

The challenges that the research library community now face have some similarities to what the computer and information science communities regarded as opportunities in the early and mid-1990s. Atkinson was writing about the control zone in an identifiable context (before the launch of Google in 1998), at a time when digital library pioneers were pursuing a vision of "tens of thousands of repositories of digital information that are autonomously managed yet integrated into what users view as a coherent digital library system" (Lynch and Garcia-Molina 1995, under Executive Summary section III).

Atkinson concluded his 2006 conference paper with the thought that "none of these challenges can be met by research libraries working independently. They can only be confronted collectively." There is reason for optimism if the research library community

collectively reexamines the early vision of thousands of interconnected repositories and identifies what elements of that vision are relevant today. Thanks to twenty years' work on digital libraries and repositories, many more tools, knowledge and expertise are available now to assemble a more coherent scholarly research infrastructure to which research libraries can substantively contribute. What is lacking are on the one hand, the collective will to begin; and on the other, a commitment to face unequivocally forward, fully embracing an approach that builds on and for the web.

5.8 Questions

The previous sections lay out one possibility among many for consideration. Whether or not the ideas sketched in the previous sections prove worthy of further thought, a collective forum or series of forums that bring stakeholders together could explore a range of new ideas for collaboration and collective action. Nine questions that might be considered in such a forum or forums follow. Some of them have been asked before or in different forms; others are new. They are offered to stimulate further research library community discussions about the future of bibliographic control, cataloguing and catalogues, cooperative cataloguing systems, and how all of these relate to digital scholarship and research libraries' missions.

5.8.1 Nine questions to consider

1. Web-based discovery and access methods generally use fully automated processing and low-barrier standards. In comparison, library cataloguing tends to be more complex. Do the benefits of library cataloguing, as it is currently practiced, still justify its additional

costs? Are there other methods, better integrated with web practices and metadata supply chains, that are already being used or that could be further developed to meet the needs of researchers, instructors and students?

2. The traditions of book cataloguing have stood apart from the practices of publishers and booksellers. Should these traditions and practices continue to exist and evolve separately? Might national libraries with responsibility for legal deposit programmes directly use metadata from other participants in the metadata supply chain for books?
3. Cooperative cataloguing systems are large stores of MARC records from which cataloguers download copies for separate local catalogues. Are there other means to achieve the same ends? What would it take to move to sharing records/holdings information at the network level as opposed to copying records for local use?
4. How might current or future cooperative cataloguing systems provide for e-content discovery and management at the network level so that investments in redundant local infrastructures and costs (e.g., for local instances of e-resource management systems and knowledge bases, duplicative purchases of MARC record sets for e-content, etc.) might be re-couped?
5. In what ways might current or future cooperative cataloguing systems provide additional incentives or more support for efficiently creating and enriching resource descriptions of unique or hidden materials?
6. In what ways might current or future cooperative systems raise awareness and stimulate more use of library print collections?

7. What roles might network-level, web services-enabled registries and/or knowledge bases (e.g., for holdings, work sets, persistent identifiers, researcher names, etc.) play in current or future cooperative library systems?
8. In what ways might today's cooperative cataloguing systems be altered to help research libraries advance their roles supporting scholarship?
9. How might a new library service framework blend with the information-seeking practices and preferences of scholars and students, attracting their attention and helping them to get their work done? What are the best ways to reach out to, engage with and include scholars and students in future research library conversations about local services, domain-specific services, regional or national services, and/or new library cooperative commons services?

Acknowledgments

This paper is a personal commentary, and all opinions and any errors are the responsibility of the author. It includes rewritten and substantially expanded portions of an earlier short paper, which grew out of a strategic analysis undertaken in 2009 with OCLC colleagues John Chapman, Ted Fons, Glenn Patton, Renee Register, Bob VanVolkenburg and David Whitehair. That early paper was translated into Japanese by Hisako Kotaka and published two years ago in a Japanese information science journal (Calhoun 2010). The author is grateful for these colleagues' contributions to the early paper and for the opportunity that Sally Chambers provided to update and rewrite it for this book.

References

Anderson, Rick. 2011. “Print on the Margins: Circulation Trends in Major Research Libraries.” *Library Journal*. http://www.libraryjournal.com/lj/home/890835-264/print_on_the_margins_circulation.html.csp.

Association of Research Libraries. 2008. “Special Collections Task Force Final Status Report, 2006.” http://www.arl.org/rtl/speccoll/spcolltf/status0706.shtml.

Atkinson, Ross. 1996. “Library Functions, Scholarly Communication, and the Foundation of the Digital Library: Laying Claim to the Control Zone.” *The Library Quarterly*: 239–265.

Atkinson, Ross. 2006. “Six Key Challenges for the Future of Collection Development.” *Library Resources & Technical Services* 50 (4): 244–251.

ANDS: Australian National Data Service (Clayton VIC: ANDS, 2012) http://www.ands.org.au/andsbrochure.pdf.

Bade, David. 2009. “Irresponsible Librarianship: a Critique of the Report of the Library of Congress Working Group on the Future of Bibliographic Control and Thoughts on How to Proceed.” http://eprints.rclis.org/bitstream/10760/12804/1/MOUG_2009.pdf.

Bell, Michael Everette. 2003. “Human Interaction at Earth’s Largest Library: The Current State of a Global Information Infrastructure Design (2003).” Unpublished manuscript. http://syndeticsystems.com/Bell%20-%20Human%20Interaction%20At%20Earth’s%20Largest%20Library.pdf.

Björk, Bo-Christer, A. Roos, and M. Lauri. 2009. “Scientific Journal Publishing–Yearly Volume and Open Access Availability.” *Information Research* 14 (1): 391.

Björk, Bo-Christer, Patrik Welling, Mikael Laakso, Peter Majlender, Turid Hedlund, and Guðni Guðnason. 2010. “Open Access to the Scientific Journal Literature: Situation 2009.” *PLoS ONE* 5 (6) (June). doi:10.1371/journal.pone.0011273.

Borgman, Christine L. 2007. *Scholarship in the Digital Age*. MIT Press.

Breeding, Marshall. 2012. “Automation Marketplace 2012: Agents of Change .” Blog entry in *The Digital Shift,* March 29. http://www.thedigitalshift.com/2012/03/ils/automation-marketplace-2012-agents-of-change/.

Burton, A., and A. Treloar. 2009. “Publish My Data: A Composition of Services from ANDS and ARCS.” In *e-Science, 2009. Fifth IEEE International Conference On*, 164–170. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5380872.

Calhoun, Karen. 2006. *The Changing Nature of the Catalog and Its Integration with Other Discovery Tools*. Washington DC: Library of Congress. http://www.loc.gov/catdir/calhoun-report-final.pdf.

Calhoun, Karen. 2010. “Rethinking bibliographic control in the Web environment.” Trans. Hisako Kotaka. *Joho no kagaku to gijutsu = Journal of Information Science and Technology Association* 60 (9): 378–383.

Calhoun, Karen. 2011. “Library Process Redesign: Renewing Services, Changing Workflows” Presented in February at Cambridge University Library. http://www.slideshare.net/amarintha/library-process-redesign-renewing-services-changing-workflows.

Chad, Ken. 2009. *Creating Catalogues: Bibliographic Records in a Networked World*. A Research Information Network Report. Research Information Network. http://rinarchive.jisc-collections.ac.uk/our-work/using-and-accessing-information-resources/creating-catalogues-bibliographic-records-network

CIBER. 2009. *E-Journals: Their Use, Value and Impact: a Research Information Network Report*. Research Information Network. http://www.rin.ac.uk/system/files/attachments/E-journals-report.pdf.

Coffman, Steve. 1999. “Building Earth’s Largest Library: Driving into the Future.” *SEARCHER* 7 (3): 34–47. http://www.fileformat.info/other/library/coffman_mar99.pdf

Cole, Christopher, Jennifer Marrill, Diane Boehr, Dianne McCutcheon, and Beacher Wiggins. 2011. *Report and Recommendations of the U.S. RDA Test Coordinating Committee*. U.S. RDA Test Coordinating Committee. http://www.loc.gov/bibliographic-future/rda/source/rdatesting-finalreport-20june2011.pdf.

Cox, John, and Laura Cox. 2008. *Scholarly publishing practice : academic journal publishers’ policies and practices in online publishing*. Brighton, UK: Association of Learned and Professional Society Publishers.

Dempsey, Lorcan. 2006. “The Library Catalogue in the New Discovery Environment: Some Thoughts.” *Ariadne* (48): 4. http://www.ariadne.ac.uk/issue48/dempsey.

Dempsey, Lorcan. 2012. “Linking not typing … knowledge organization at the network level.” Blog entry in *Lorcan Dempsey’s Weblog*, January 1. http://orweblog.oclc.org/archives/002195.html.

De Rosa, Cathy. 2005. *Perceptions of libraries and information resources : a report to the OCLC membership*. Dublin, Ohio: OCLC Online Computer Library Center. http://www.oclc.org/reports/2005perceptions.htm.

De Rosa, Cathy, Joanne Cantrell, Matthew Carlson, and Peggy Gallagher. 2011. *Perceptions of Libraries, 2010: Context and Community*. Report to the OCLC Membership. Dublin

OH: OCLC Online Computer Library Center. http://www.oclc.org/reports/2010perceptions.htm.

Dooley, Jackie M, and Katherine Luce. 2010. *Taking our pulse: the OCLC Research survey of special collections and archives*. Dublin, Ohio: OCLC Research. http://www.oclc.org/research/publications/library/2010/2010-11.pdf.

Duranceau, Ellen Finnie, and Richard Rodgers. 2010. "Automated IR Deposit via the SWORD Protocol: An MIT/BioMed Central Experiment." *Serials: The Journal for the Serials Community* 23 (3): 212–214. http://uksg.metapress.com/content/l437x1631052407r/fulltext.pdf.

Eden, Brad. 2011. "The Status Quo Has Got To Go." Presented at the Charleston Conference, November 8. http://www.slideshare.net/CharlestonConference/the-status-quo-has-got-to-go-by-brad-eden-dean-of-library-services-valparaiso-university-sat-930-am.

El-Sherbini, M., and A. J. Wilson. 2007. "New Strategies for Delivering Library Resources to Users: Rethinking the Mechanisms in Which Libraries Are Processing and Delivering Bibliographic Records." *The Journal of Academic Librarianship* 33 (2): 228–242.

Fischer, Ruth, and Rick Lugg. 2009. *Study of the North American MARC Records Marketplace*. R2 Consulting. http://www.loc.gov/bibliographic-future/news/MARC_Record_Marketplace_2009-10.pdf.

Gardner, Sue Ann. 2012. "Cresting Toward the Sea Change." *Library Resources & Technical Services* 56 (2): 64.

Green, Richard, and Chris Awre. 2009. "Towards a Repository-enabled Scholar's Workbench." *D-Lib Magazine* 15 (5/6) (May). doi:10.1045/may2009-green. http://dlib.org/dlib/may09/green/05green.html.

Guenther, Rebecca, and Jacqueline Radebaugh. 2004. *Understanding metadata*. Bethesda, Md.: NISO Press. http://www.niso.org/publications/press/UnderstandingMetadata.pdf.

Hampton-Reeves, S., C. Mashiter, J. Westaway, P. Lumsden, H. Day, and H. Hewertson. 2009. *Students' Use of Research Content in Teaching and Learning: a Report for the Joint Information Systems Council (JISC)*. University of Central Lancashire, Centre for Research-Informed Teaching. http://ie-repository.jisc.ac.uk/407/1/Students_Use_of_Research_Content.pdf.

Henry, Charles, Lisa Spiro, Geneva Henry, Paul Courant, Matthew Nielsen, Kathlin Smith, and Roger C. Schonfeld. 2011. *The Idea of Order : Transforming Research Collections for 21st Century Scholarship.* CLIR Publication. Washington DC: Council on Library and Information Resources. http://www.clir.org/pubs/reports/pub147/reports/pub147/pub147.pdf.

Hruska, Martha. 2009. “Where Are We With the Staffing Transition From Cataloging to Metadata Management?” *Technicalities* 29 (2) (April 1): 1.

Hruska, Martha. 2011. “Is It the Next Generation Yet?” *Technicalities* 31 (4) (July 1): 1.

Laakso, Mikael, Patrik Welling, Helena Bukvova, Linus Nyman, Bo-Christer Björk, and Turid Hedlund. 2011. “The Development of Open Access Journal Publishing from 1993 to 2009.” *PLoS ONE* 6 (6). doi:10.1371/journal.pone.0020961.

Lagoze, Carl, Herbert Van de Sompel, Michael L. Nelson, Simeon Warner, Robert Sanderson, and Pete Johnston. 2008. “Object Re-Use & Exchange: A Resource-Centric Approach.” *arXiv:0804.2273* (April 14). http://arxiv.org/abs/0804.2273.

Lavoie, Brian, L. S. Connaway, and E. T. O’Neill. 2007. “Mapping WorldCat’s Digital Landscape.” *Library Resources & Technical Services* 51 (2): 106–115.

Lavoie, Brian, Geneva Henry, and Lorcan Dempsey. 2006. “A Service Framework for Libraries.” *D-Lib Magazine* 12 (7/8) (July). doi:10.1045/july2006-lavoie. http://www.dlib.org/dlib/july06/lavoie/07lavoie.html.

Lewis, David W. 2013. “From Stacks to the Web: The Transformation of Academic Library Collecting.” *College & Research Libraries* (forthcoming).

Lewis, Stuart, Pablo de Castro, and Richard Jones. 2012. “SWORD: Facilitating Deposit Scenarios.” *D-Lib Magazine* 18 (1/2) (January). doi:10.1045/january2012-lewis. http://www.dlib.org/dlib/january12/lewis/01lewis.html.

Library of Congress. 2012a. “Bibliographic Framework Transition Initiative.” http://www.loc.gov/marc/transition/.

Library of Congress. 2012b. “The Library of Congress Announces Modeling Initiative (May 22, 2012): Bibliographic Framework Transition Initiative (Library of Congress).” http://www.loc.gov/marc/transition/news/modeling-052212.html.

Library of Congress. Working Group on the Future of Bibliographic Control. 2008. *On the record: report of the Library of Congress Working Group on the Future of Bibliographic Control.* Washington DC: Library of Congress. http://www.loc.gov/bibliographic-future/news/lcwg-ontherecord-jan08-final.pdf.

Loughborough University. Library & Information Statistics Unit, and Research Information Network. 2007. *Uncovering hidden resources : progress in extending the coverage of online catalogues: report of a study undertaken for the Research Information Network*. [London]: Research Information Network. http://rinarchive.jisc-collections.ac.uk/our-work/using-and-accessing-information-resources/uncovering-hidden-resources-extending-coverage-on.

Luther, M. Judy, National Information Standards Organization (U.S.), and OCLC. 2009. *Streamlining book metadata workflow*. Baltimore, MD; Dublin, Ohio: NISO ; OCLC, Inc. http://www.niso.org/publications/white_papers/StreamlineBookMetadataWorkflowWhitePaper.pdf.

Lynch, Clifford, and Hector Garcia-Molina. 1995. *Interoperability, Scaling, and the Digital Libraries Research Agenda: a Report on the May 18-19, 1995 IITA Digital Libraries Workshop*. http://www.sis.pitt.edu/~repwkshop/papers/DL1995.pdf.

Malik, Shiv. 2012. "Kindle Ebook Sales Have Overtaken Amazon Print Sales, Says Book Seller." *The Guardian*, August 5. http://www.guardian.co.uk/books/2012/aug/06/amazon-kindle-ebook-sales-overtake-print.

Malpas, C. 2011. *Cloud-sourcing Research Collections: Managing Print in the Mass-digitized Library Environment*. Dublin OH: OCLC, Inc. http://dcommon.bu.edu/xmlui/handle/2144/1391.

Michener, William, Dave Vieglais, Todd Vision, John Kunze, Patricia Cruse, and Greg Janée. 2011. "DataONE: Data Observation Network for Earth — Preserving Data and Enabling Innovation in the Biological and Environmental Sciences." *D-Lib Magazine* 17 (1/2) (January). doi:10.1045/january2011-michener

Napier, Mark E., and Kathleen A. Smith. 2000. *Earths Largest Library -- Panacea or Anathema? A Socio-Technical Analysis*. Bloomington IN: Center for Social Informatics. Indiana University. https://scholarworks.iu.edu/dspace/handle/2022/165.

Nitecki, Danuta, Carol Jones, and Jeffrey Barnett. 2009. "Borrow Direct: a Decade of a Sustained Quality Book-lending Service." *Interlending & Document Supply* 37 (4) (November 13): 192–198. doi:10.1108/02641610911006283.

Norris, M., C. Oppenheim, and F. Rowland. 2008. "Finding Open Access Articles Using Google, Google Scholar, OAIster and OpenDOAR." *Online Information Review* 32 (6): 709–715.

Payne, L. 2007. *Library Storage Facilities and the Future of Print Collections in North America*. Dublin OH: OCLC Programs and Research. http://csul.fcla.edu/sites/csul.fcla.edu/uploads/ss-oclc-storage-rept-11-02-07.pdf.

Reynolds, Regina Romano, Bruce Knarr, Library of Congress. On the Record Report Implementation Working Group, and Library of Congress. Library Services. 2009. *On the record: report Recommendations the Library of Congress should pursue over the next four years report to the associate librarian for library services*. http://bibpurl.oclc.org/web/38884 http://www.loc.gov/bibliographic-future/news/OTR_rep_response_final_091509.pdf.

Schonfeld, Roger. C., and K. M. Guthrie. 2007. “The Changing Information Services Needs of Faculty.” *EDUCAUSE Review* 42 (4): 8.

Schonfeld, Roger C, and Ross Housewright. 2010. *Faculty survey 2009: key strategic insights for libraries, publishers, and societies*. Ithaka. http://bibpurl.oclc.org/web/40082

Schuitema, Joan E. 2010. “The Future of Cooperative Cataloging: Curve, Fork, or Impasse?” *Cataloging & Classification Quarterly* 48 (2-3) (February 10): 258–270. doi:10.1080/01639370903536088.

Sellberg, R. 2010. “Cooperative Cataloging in a Post-OPAC World.” *Cataloging & Classification Quarterly* 48 (2-3): 237–246.

Soehner, C., C. Steeves, and J. Ward. 2010. *E-Science and Data Support Services: A Study of ARL Member Institutions*. Washington DC: Association of Research Libraries.

University Leadership Council. 2011. *Redefining the Academic Library: Managing the Migration to Digital Information Services*. Washington DC: Education Advisory Board. http://www.educationadvisoryboard.com/pdf/23634-EAB-Redefining-the-Academic-Library.pdf.

Van de Sompel, Herbert, Sandy Payette, John Erickson, Carl Lagoze, and Simeon Warner. 2004. “Rethinking Scholarly Communication.” *D-Lib Magazine* 10 (9) (September). doi:10.1045/september2004-vandesompel. http://www.dlib.org/dlib/september04/vandesompel/09vandesompel.html.

Wilson, Amanda J. 2007. “Toward Releasing the Metadata Bottleneck.” *Library Resources & Technical Services* 51 (1): 16–28.

Witt, Michael. 2010. “Chapter 4: Implementations of ORE.” *Library Technology Reports* 46 (4): 26–34.

Wolven, Robert. 2009. “Cataloging Without Silos, or, Where Do We Go From Here?” *Technicalities* 29 (5) (September/October): 1, 6-9.

Zillman, Marcus P. 2012. “Deep Web Research 2012.” http://www.deepwebresearch.info/