The Quest for Electronic Resource Management Standards and Tools

Adam Chandler

Our Challenge: Managing Electronic Resources

Libraries are experiencing a monumental transformation away from the traditional practice of purchasing print materials, and towards acquiring access to electronic resources. The Association for Research Libraries examined this trend in a recent report, estimating that "the percentage of the average library budget that is spent on electronic materials has increased more than eightfold, from an estimated 3.6% in 1992–93 to 25% in 2002-03."[1] Most agree that print will not disappear entirely, especially for books. However, it is now clear that databases and electronic journals are heavily preferred by patrons for many research needs. A 2002 survey conducted by Outsell found that 90% of students and faculty said they looked online first, and then consulted print sources later if needed.[2] The question is: how do we integrate this change into existing library workflows and systems?

Libraries are now undergoing a change comparable to what happened, beginning in the 1970s, as libraries made the switch to online public access catalogs. As an Information Technology Librarian working in a large technical services department (78 FTE), I have experienced first-hand the difficulties of navigating this rapidly changing environment. The core tool that libraries use to manage collections—the integrated

---

[1] Mark Young and Martha Kyrillidou, *ARL Supplementary Statistics* 2002-03 (Washington, DC: Association of Research Libraries, 2004), http://www.arl.org/bm~doc/sup03.pdf (8 Mar. 2008).
[2] Scott Carlson, "Students and Faculty Members Turn to Online Library Materials Before Printed Ones, Study Finds," *Chronicle of Higher Education* (October 3, 2002), http://chronicle.com/free/2002/10/2002100301t.htm (8 Mar. 2008).

library system (ILS)—originated in the 1970s and '80s, and was originally optimized for print materials. The current generation's challenge is to build a new kind of library system optimized for electronic resources.

Administrative metadata is essential to meeting this challenge. In the book *Metadata Fundamentals*, Priscilla Caplan distinguishes administrative metadata from descriptive metadata in the following passage:

While descriptive metadata is intended to help in finding, discovering, and identifying an information resource, administrative metadata is intended to facilitate the management of the resource. Management functions typically include such activities as tracking an item through various stages of processing, controlling access to the resource, establishing responsibilities related to the resource, and granting permission for its use. Although this is overly simplistic, descriptive metadata can be thought of as serving the actual or potential users of a resource, while administrative metadata serves the owners or caretakers of the resource.[3]

Electronic resources, especially those that are accessed remotely for a fee, require administrative metadata that must be maintained over the entire lifecycle of the resource. Examples of administrative metadata include: lists of IP addresses that are authorized for access to a database, information about whether or not interlibrary loan of the contents is allowed, and the user name and password required to access the vendor's administrative module. Many of these elements were not relevant in the print-only world, but are now critical in today's library.

This essay is divided into two sections. The first covers a national effort by librarians to articulate the functional requirements and develop standards for electronic resource management. The second section describes highlights of Cornell University Library's (CUL) effort since 1999 to gain control over its administrative metadata for electronic resources.

---

[3] Priscilla Caplan, *Metadata Fundamentals for All Librarians* (Chicago: American Library Association, 2003), 151.

A Brief History of the Electronic Resource Management Initiative Project

My involvement in the standardization of administrative metadata for electronic resources started in the fall of 2000 when Karen Calhoun, my supervisor and also Associate University Librarian for Technical Services here at Cornell University Library, asked me to explore the possibility of building a database system to help manage local electronic resource licensing metadata. When given this assignment, I intended to do some research and start constructing a system that could be used at Cornell, leaving open the possibility of making it available for use in other libraries. To begin, I described the project in an application for the American Library Association's Samuel Lazerow Fellowship.[4] The same month, I learned that Timothy Jewell, Head of Collection Management services at the University of Washington, was pursuing a similar research thread. Jewell was gathering information and making contact with pioneers at UCLA, Johns Hopkins University, Yale University and elsewhere who had built (or planned to build) local electronic resource management systems. Tim and I communicated over email and telephone and quickly agreed to collaborate. In early 2001, I began working on a Web site that would bring together the information that Jewell and I had uncovered. The Web site became known as the "Web Hub," and it would help to connect people who are working on the same problem.[5] In the summer of 2001, Jewell's report on his survey, "Selection and Presentation of Commercially Available Electronic Resources: Issues and Practices," was published by the Digital Library Federation (DLF)

---

[4] Adam Chandler, "An Application Profile and Prototype Metadata Management System for Licensed Electronic Resources." Proposal for 2001 Samuel Lazerow Fellowship for Research in Collections and Technical Services in Academic and Research Libraries, Association of College and Research Libraries, American Library Association, December 2000.

[5] Norm Medeiros, "A Pioneering Spirit: Using Administrative Metadata to Manage Electronic Resources." *OCLC Systems & Services* 19, no. 3 (2003).

and the Council on Library and Information Resources (CLIR).[6] This seminal document defines the most important management issues facing libraries as serials content migrates from print to digital, and from ownership to access.

In the course of writing the DLF/CLIR report, Jewell developed a working relationship with Daniel Greenstein (Director at the time of the DLF). Greenstein was instrumental in helping Jewel bring together other librarians and vendors as well as Pat Harris of the National Information Standards Organization (NISO) for the half-day pre-standardization investigation held on May 10, 2002, entitled "NISO/DLF Workshop on Standards for Electronic Resource Management." The Workshop broke new ground and made it clear that the community was ready to place sustained, focused attention on the problem of standardizing the management of electronic resources. The organizing committee for the workshop later became the Steering Group for the DLF Electronic Resource Management Initiative (ERMI).[7] The Steering Group was comprised of the following seven librarians: Ivy Anderson (Harvard), Sharon Farb (UCLA), Kimberly Parker (Yale), Angela Riggio (UCLA), and Nathan Robertson (Johns Hopkins), plus Jewell and myself.

The primary product of the DLF ERMI Steering Group was a series of documents published on the Web in August 2004, "Electronic Resource Management: The Report of the DLF Initiative".[7] After the May 2002 DLF/NISO Workshop ended, the steering group discussed its effectiveness. One problem we observed was the difficulty in communicating to others which metadata elements are needed. As a response, we

[6] Timothy D. Jewell, "Selection and Presentation of Commercially Available Electronic Resources: Issues and Practices," (CLIR & DLF, 2001), http://www.clir.org/pubs/reports/pub99/contents.html (8 Mar. 2008).
[7] Timothy D. Jewell, Ivy Anderson, Adam Chandler, Sharon E. Farb, Kimberly Parker, Angela Riggio, and Nathan D. M. Robertson, "Electronic Resource Management: The Report of the DLF Initiative," (Washington, DC: Digital Library Federation, August 2004), http://www.diglib.org/pubs/dlf102/ (8 Mar. 2008).

developed an entity-relationship diagram that defines a place for each element in relation to all other elements. This work is presented in "Appendix C: Entity Relationship Diagram for Electronic Resource Management." An alphabetical list of all the elements is contained in "Appendix D: Data Element Dictionary" followed by a more complete data model in "Appendix E: Electronic Resources Management System Data Structure." Jewell authored the "Final Report of the DLF Initiative," which surveys how various locally-built systems attempt to manage electronic resources. "Appendix A: Functional Requirements" started as a document that Ivy Anderson and Ellen Duranceau (MIT) wrote to guide the creation of an "ERM" product from *Ex Libris*, the library systems company. Anderson and the other members of the DLF ERMI steering group then derived a more general document that could be applied across many institutions. In the course of discussions about the workflow of electronic resource management, we felt it would be useful to create a detailed model of an institution's expected workflow—since releasing the report, we have heard vendors, including Endeavor and VTLS, have said that "Appendix B: Electronic Resource Management Workflow Flowchart" has been extremely valuable in their software design process.

Since release of the report in August 2004, the topic that has perhaps generated the most controversy is how to express and communicate DLF ERMI license terms in XML. In this debate, the stakes are high. The esoteric characteristics of the rights expression language that emerges as the standard for the exchange of license terms between the various parties (authors, publishers, aggregators, libraries, users) will define legitimate "fair use." I served as primary author on "Appendix F: XML Investigation," where this issue is discussed in detail. There are many outstanding questions in this realm: Is the XML markup a one-to-one representation of the actual

license? Exactly which DLF ERMI elements need to be represented in XML for exchange? Is it possible to place the elements in an XML wrapper that moves along from the publisher source to the end user? As a result of the dialogue, there is a subtle change happening in the discourse among some librarians: "rights expression" is changing to "license expression," as attention is being focused on a relatively narrow library-oriented use case that involves the exchange of standardized license terminology. For more detail, please see Appendix F of the DLF ERMI report.

Case Study: Electronic Resource Management at Cornell

Cooperative Online Research Cataloging

In 1999, a cross-functional team of Cornell Library staff was formed to investigate OCLC's Cooperative Online Research Cataloging (CORC) system. CORC was an experimental project, developed between 1998 and 1999, that evolved into OCLC's Connexion metadata management product. According to OCLC:

> CORC was one of OCLC's most exciting and ambitious projects. CORC explored the cooperative creation by libraries of a database of Web resources. Features of the prototype system included:
> 1. Automation-assisted generation of metadata with DDC numbers,
> 2. Tools for pathfinder page generation and editing,
> 3. Automation-assisted URL maintenance,
> 4. Loading and output of several standard forms of metadata, and
> 5. A host of standard database and information system functions, including relevance-ranked retrieval and Unicode-support for searching and display. CORC transitioned from research project to OCLC service in 1999.[8]

The thrust of the CORC project was *descriptive* metadata. The Cornell team concluded that the CORC product, although interesting, lacked a key piece of the puzzle, the capacity to record and maintain *administrative* metadata for electronic resources.

[8] Online Computer Library Center (OCLC), Cooperative Online Resource Catalog (CORC), http://www.oclc.org/research/projects/archive/default.htm (8 Mar. 2008).

An appendix to the Cornell report lists an array of elements that would eventually find their way into the DLF ERMI report, such as the pricing, number of simultaneous users, access restrictions, and hyperlinks to the full text of the license agreement.[9] One could speculate that OCLC missed an early opportunity to define and bring to market an ERM product.

As stated earlier, my original intention was to build a stand-alone electronic resource management system for Cornell. Other libraries, such as Johns Hopkins and UCLA, followed that path around the same time. The more I learned about the problem, however, the more I felt it would be best to invest in promoting national standards so that other parties, such as library systems developers, could build systems for many libraries to use. Turnkey systems would come with a service contract.

Choosing the Innovative Interfaces Inc. ERM stand-alone module

Cornell uses a number of Endeavor products, including EnCompass and Voyager, so naturally we considered using the ERM product being developed by Endeavor, but they were not ready with a product in time. By 2004, Innovative Interfaces, Inc. (III) was in production with an ERM module integrated into their library management system. After comparing the III product to our internal requirements (which were derived from the DLF ERMI Appendix A: Functional Requirements document), CUL signed a contract in the summer of 2004 with III to implement the stand-alone version of its ERM module. Other early customers of III's ERM were the Library of Congress, Utah State University, and more recently, Stanford University. A cross-functional team of librarians implemented the III system at Cornell in March 2005, beginning as a

_____

[9] Karen Calhoun, et al., "CORC at Cornell Project: Final Report: Appendix 1," (December 1999), http://www.library.cornell.edu/corc/ (8 Mar. 2008).

replacement for the Endeavor ENCompass-based "Find e-Journals" service. The challenge of integrating a product from a different vendor, of course, is mapping the data, and keeping it in sync between the different systems, i.e. Endeavor's Voyager library management system and III's Electronic Resource Management module. Eliminating redundant data management was one of the key functional requirements within the DLF ERMI model.
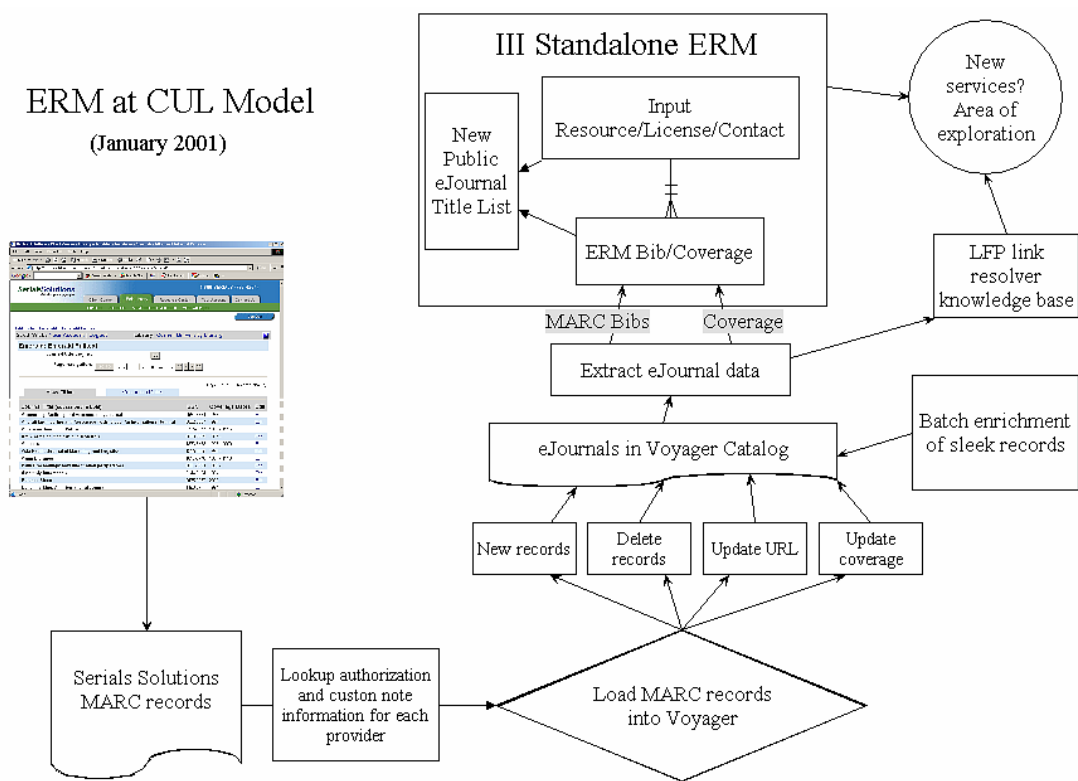

Figure 7.1 Cornell Electronic Journal Management Model, April 2005

Figure 7.1 is a representation of the way electronic journals are being managed by Cornell in April 2005. We have learned that electronic journals are a moving target: this figure describes what we are doing right now, but this model will most definitely evolve.

There are several points to highlight. First, coverage data (title, ISSN if available, provider, start date, end date, embargo period) are being managed for each electronic journal aggregator (e.g. JSTOR, Proquest, and Wiley) within the Serials Solutions vendor interface. Each month Serials Solutions makes available to Cornell the most recent full MARC records for our titles. Second, after a lookup is made for each provider to see what kind of authorization is allowed for the set, those records are loaded into the Voyager catalog. Special keys within the 899a field are inserted into the Voyager bibliographic record which then links it to the ERM "resource record" when the MARC records are later exported and loaded into the III ERM Module. Finally, once that relationship is made, it is possible to associate license terms with the appropriate titles. The ERM maintains all the CUL license terms and contact information.

Conclusion

Today, tools are being developed that are based on the model put forward by the DLF ERMI steering group. All of the major library software companies, Visionary Technology in Library Solutions (VTLS), Dynix, Endeavor, and Ex Libris—to name a few—are moving forward with products. On the other side of the fence, a small number of libraries, such as Boston College and North Carolina State, are using the DLF ERMI model guide to develop their own second generation systems. Over the next two to four years ERM software will be integrated into library workflows everywhere.

The primary objective for an ERM system is to improve the quality of service for library users, by providing up-to-date information about the library's electronic resources. As the library establishes better control over electronic resources, users will have access to more accurate information about those resources, even as titles are added

or dropped by service providers, or as date coverages are updated. Although there is no empirical evidence yet that proves libraries can serve patrons better with the advent of ERM technology, I am confident that future studies will show this to be true. The Cornell experience indicates that ERMs enable libraries to manage, in one place, the complex administrative metadata that is required over the lifecycle of the library's rapidly expanding collection of electronic resources.