

House of Horrors: Exorcising Electronic Resources

Norm Medeiros

Associate Librarian of the College
Haverford College
Haverford, PA
nmedeiro@haverford.edu

A published version of this paper appears in *Managing Electronic Resources: Contemporary Problems and Emerging Issues* (Chicago: ALA, 2006)

INTRODUCTION

Electronic resources, especially electronic journals, are maddening. Come to think of it, print serials have never been much fun to deal with either. C. Sumner Spalding, writing in 1957 and in a tone appropriate for the science fiction craze of the time, summarized well the frustrations serials catalogers experience (Spalding, 1957):

"Whereas monographs are normally received in their fully-developed form, physically static, ossified, and dead, so to speak, serials are generally received alive, growing, and subject to various unpredictable metamorphoses. The problem of describing something which is alive and which may at any time in the future assume relationships with others of its kind ... is of a different order from that of describing something which can be counted on to lie still and not move while we take its measure and describe it[s] characteristics once and for all."

Thirty years after Spalding's remarks, Syracuse University's *New Horizons in Adult Education* electronic journal debuted, the first of a prolific species that often make me feel nostalgic about my Kardex days.

To best understand today's crazy world of electronic resources, we need to reflect on the evolution of the e-journal. Like Spalding, another editor of the *Anglo-American Cataloging Rules*, Michael Gorman, published in 1991 an article in which he made certain predictions about libraries in the coming decade. Not surprisingly, many of his predictions came true. One remarkably incorrect predication, however, was that the impact of electronic journals would "rise from the invisible to the miniscule in the next five years," and will be "unimportant" in 2001 (Gorman, 2001). You really can't blame him for miscalculating the impact electronic publishing would have on libraries. When he wrote this article, there were eight peer-reviewed electronic journals in existence, produced by pioneering academics at institutions such as Virginia Tech, Princeton, and the aforementioned Syracuse University. Commercial publishers were nowhere to be seen. Today, 75% of the scholarly journals published offer online access, with few, if any, commercial publishers not in the game (Cox, 2003).

These early e-journals were sent to subscribers as issues appeared. Occasionally, a subscriber would receive notification that an issue was ready, and she would send a

command to a listserv in order to receive the full text. As early as 1990, libraries began worrying about archival access to these publications, many storing local copies of e-journals electronically or in print (McMillan, 1993). Since the publishers of these early e-journals were academics, there were no guarantees regarding length of publication. If an editor decided to stop issuing the e-journal, or he moved to another university, there was no telling what might happen to the scholarship he had overseen. Naturally, there was great concern over the transient nature of this new publishing process, not to mention questions about how libraries would make these journals available to their patrons. Remember, these were the pre-Web days, when library staff used dumb terminals to access OCLC and their mainframe-based library systems. BITNET and fledgling Internet email were recent innovations, and text-based Gopher servers were appearing. There were no standards for cataloging electronic journals, and certainly no best practice guidelines to help libraries provide access to these resources.

Despite these problems, there was in hindsight much good fortune, starting with the small number of available titles. As noted above, there were eight peer-reviewed e-journals available in 1991, and still only about 100 at the start of 1995 (Hitchcock, 1996). Serialists in the early 1990s also benefited from knowing what had been published. Subscribers were notified of new issues. Action could be taken to process these issues. At Virginia Tech for example, electronic journals followed the same procedures as print serials (McMillan, 1993). Issues were checked in, and claimed as necessary. Hard to imagine doing this today given the hundreds or thousands of e-journals licensed by most academic libraries. Furthermore, demand for these early e-journals was low, so libraries were able to experiment with procedures and workflows. It was truly a grace period that ended when the Mosaic browser illuminated the World Wide Web.

WEB REVOLUTION

Mosaic was released in 1993. It was the first widely-used graphical browser, and within 18 months helped bring about the Web revolution. On the heels of this innovation was the implementation of OCLC's Internet Cataloging Project in 1994. The following year, the MARC 856 field was approved, which turned the OPAC from atlas to vehicle of information delivery, and spawned rapid development by opac vendors of web-enabled catalogs. Also in 1995, the first edition of Nancy Olson's seminal guide to cataloging internet resources was published. Amidst the innumerable changes resulting from Mosaic and the Web, publishers began tinkering with online access to their journals. It seems ridiculous in hindsight, but at the time no one knew if print journals delivered via the web would be popular or marketable. In some part because of this uncertainty, many publishers gave away their online content to institutions that were print subscribers. Libraries spent hours determining how to catalog these resources. Fledgling web developers put up annotated e-journal lists, and not remarkably, people used them. Some technically competent professors – at that time defined as those who used email regularly – took advantage of the new offerings. Libraries didn't mind providing the service, since the resources were free, and although it meant cataloger and web personnel staff time, it seemed worth it. Many academic libraries in the mid-1990s in fact were leaders on their campuses when it came to having a web presence.

Development of e-journal lists was among the first value-added content to grace these pages.

Almost overnight, providing access to electronic journals was seen as one of the most important aspects of technical services. Jim Holmes, Head of Serials Cataloging at the University of Texas at Austin, described the impact well in his 1998 article that appeared in *Serials Librarian* (Holmes, 1998). "Some time late in 1995, an email message arrived in technical services from the Associate Director for Technical Services. She was so excited about URLs and told technical services that it should also be excited and should be giving URLs a lot of thought. About 100 fingers pressed the delete button, trashed the message, and got back to cataloging the *Newsletter of the Texas Republic*." Jim goes on to say that within months, UT Austin had purchased access to JSTOR, Project Muse, and the American Institute of Physics collections, representing over 6,000 URLs, and the old adage about job security was not comforting.

ACCESS VERSUS OWNERSHIP

During this period, ideological debates about the role of the catalog raged. Should it be an inventory list or a finding aid?. Regina Reynolds, Head of the National Serials Data Program at the Library of Congress, wrote an especially compelling piece about this issue in the Winter 1995 edition of *Serials Review* (Reynolds, 1995). Reynolds questioned the value of pointing to remote files through the catalog, suggesting the practice could mislead users. She admitted, however, the transformation of the catalog from inventory list to information gateway was inevitable. Despite this debate, many libraries began getting records for e-resources into their catalogs, and exploiting web-enabled opacs. By the mid 1990s, many libraries provided redundant access to e-journals, both through the catalog as well as through hand-coded web pages. Inclusion of bibliographic records for e-resources in OCLC eased some of the burden, but even these records needed local adjustment to account for mode of access and other differences between the library that provided the original bibliographic record and the library cribbing from it.

In 1996, CONSER approved a single-record approach to cataloging e-journals also held in print, which once again eased some of the cataloging work, but has since proven problematic for some libraries that receive the same e-journal from two or more providers, as is the case for instance with many JSTOR, Project Muse, and other aggregator titles. Some libraries that share an online catalog, like the Tri-College Consortium of Bryn Mawr, Haverford, and Swarthmore Colleges, decided from the outset to catalog using separate records for each manifestation of an e-journal. This decision is more labor intensive, but better than the alternative that would in our case make a mockery of the overburdened bibliographic record.

Over the past few years, libraries have automated the record-creation process. Brief bibliographic records can be derived using title lists provided by e-journal aggregators, such as Expanded Academic Index. The derived brief records can be batch loaded into the catalog, resulting in quick-n-dirty opac access to these titles. Acquisition of more descriptive MARC record sets is also becoming a popular way to protect precious staff time. Automated ways of getting e-journal lists to library web sites have also been

developed over the past few years, in partial response to the debate that persists, which access point is better – the online catalog or the web site? By the end of the 1990s, database-driven web lists of e-journals replaced manually-maintained pages. These databases can not only create dynamically-generated web pages, but by using an application like MyLibrary, developed by Eric Lease Morgan, then at North Carolina State University, these systems can also create personalized web lists. Both opac and web site access to e-journals are valuable services provided by libraries. Link servers such as SFX, and course management software, further exploit this medium.

ELECTRONIC RESOURCE ADMINISTRATION

Since libraries have developed useful protocols for providing e-resource access to their users, many libraries have begun to focus on management of administrative metadata about these e-resources. What is administrative metadata? Administrative metadata are elements about licensed electronic resources that often get filed away in a cabinet or an email mailbox, or more often than not, misplaced or even discarded. They are the terms of a license, the name and contact information for a vendor representative, the purchase arrangement, and the URL where usage statistics are located. Administrative metadata are all of these elements and more, and libraries are learning just how important access to these data can be.

Libraries have struggled to organize and disseminate to staff administrative elements associated with licensed resources. Even in the golden days, when e-journals were all free, when we didn't have to create elaborate databases to manage them or lease nuclear-powered printers to reproduce their contents, some librarians knew a day would come when the access components – the ways we get users to electronic content – would mature and become second-nature, but that contracts, and the lawyers who write them, would make our lives hell forever (Moothart, 1996). As collections migrate from print to digital, and permissions from those allowed through copyright to those dictated by a contract, more and more we need to know the terms of our e-resources before we can use them.

Electronic resources, especially electronic serials, offer many of the challenges faced with print materials, but also a host of new ones. Libraries of all types and sizes purchase access to an increasing number of e-resources. These resources possess an array of restrictions, elements that don't easily fit, and were never intended, for capture in integrated library systems. Administrative metadata are increasingly needed by technical services staff, not to mention interlibrary loan and reference librarians, since use of electronic resources is governed by contract law, which is unique to each publisher. Personnel responsible for reporting to statistics-collecting agencies also need these data. For instance, one of the questions I'm asked annually is how much money the library spends on electronic resources. What seems like a straightforward question is quite complicated. For example, if Haverford purchases a journal in print for \$1,000, and electronic access to it costs an additional \$500, when this payment is posted in my library system, it looks as though we spent \$1,500 to pay for a print journal, and had electronic access granted as a result. My library system has no way of distinguishing between the amount spent for the print and the amount spent for the electronic access,

but the difference is very important when you consider the increasing number of journal publishers emigrating from the free-with-print pricing model.

The elements important to managing electronic resources, beyond providing access and what is captured within an integrated library system, generally fall into three categories: *licensing*, *purchasing*, and *administration*. Licensing involves a myriad of terms that dictate how and by whom the resource may be used. Some of these elements include restrictions on interlibrary loan, remote access, use in course management tools, concurrent users, archival access guarantees, and indemnification. Purchasing data include elements such as vendor name, expiration date, pricing model, and consortial arrangement. Administration data include elements such as OpenURL compliance, availability and frequency of usage statistics, administrative passwords and documentation, and technical contact information. These are broadly defined categories, and the elements noted above are but a sampling.

As early as 1996, librarians recognized that managing licensing agreements would be more problematic than creating and maintaining web links (Moothart, 1996). Few libraries did anything to resolve this issue back then, but shortly thereafter a handful of research libraries began building systems to help them manage administrative metadata. Some of these systems married the organization of administrative metadata with the delivery of e-resources to end-users. In this regard, these systems were unique, for although several libraries had created database-driven e-resource delivery systems by the end of the 1990s, only a few had created the administrative piece. Three early systems that have influenced later designs, as well as the standards currently under development, are MIT's Vera <<http://libraries.mit.edu/vera>>, Penn State's ERLIC <http://www.libraries.psu.edu/tas/fiscal_data/ERLIC_SHARE/Publish/index.html>, and Johns Hopkins HERMES <<http://hermes.mse.jhu.edu:8008/hermesdocs/>> systems.

I learned of Vera, ERLIC, and HERMES by sheer luck. In early 2001 I was returning by train from Washington DC while reading an article in *Library Collections, Acquisitions, and Technical Services*. The topic of the article was a database solution to managing e-journals. I was interested in the article because I had recently been charged by the Tri-College Consortium with getting a handle on our e-journal collections. The piece that needed the most work in my opinion was the licensing part, and the article discussed Griffith University's Electronic Resources Database, the type of resource I felt we needed in the Tri-Colleges (Schultz, 2001). In her article, Schultz mentioned Cornell's "Web Hub for Developing Administrative Metadata for Electronic Resource Management" <<http://www.library.cornell.edu/cts/elicensestudy/home.html>> - which was responsible for opening my eyes to the efforts of MIT, Penn State, Johns Hopkins, and other libraries - efforts I previously had no idea existed. How comforting to know that others had faced the challenge we were experiencing.

ERTS: THE ELECTRONIC RESOURCES TRACKING SYSTEM

The Tri-College Consortium consists of Bryn Mawr, Haverford, and Swarthmore colleges, three small liberal arts institutions located within a 10-mile radius in the

suburbs of Philadelphia. We've shared an integrated library system, "Tripod," an Innovative Interfaces product since 1990, and our collaborative efforts during this time have steadily grown, especially with respect to electronic resource purchasing. These efforts are often time-consuming and challenging to say the least, but they're worth it, for they allow us to operate as though we were a larger institution, both with respect to staff and budgets. Electronic-resource acquisitions within Haverford's small consortium are growing exponentially. Through the outstanding efforts of the Tri-College cataloging departments, the consortium is able to provide opac access to these e-resources, even those titles available via aggregator services. Our library system, however, does not provide a means of recording the administrative elements associated with licensed resources. Traditionally, these data have been kept in paper files or as email messages, which makes these elements difficult to resurrect when questions arise. Furthermore, our library system does not provide a mechanism for generating useful statistics about our electronic holdings. For instance, there is no way to tell how many e-journals we receive as a consequence of print subscriptions, versus how many we receive by paying an additional fee beyond the cost of the print. Prior to creating the Electronic Resources Tracking System (ERTS), we struggled with knowing when resources were up for renewal. Many expired at the end of the calendar year, while others expired at the end of the fiscal year. Still others had different expiration dates. We wanted to have a means of knowing when these e-resource subscriptions would expire before receiving a call from a faculty member wondering why the *Journal of Biological Chemistry* was asking him for a username and password.

ERTS development had its beginnings in early 2001, when I facilitated a discussion about the state of e-journal practice within the Consortium, this discussion falling under the jurisdiction of a Mellon Foundation grant we were enjoying at the time. Following a few months where the project I suggested floundered due to its completely impractical nature, a colleague walked in my office and listened to me imagine an e-resource utopia, where we had access to licensing terms at the click of a button, where I could store URLs and passwords to the usage statistics I coveted, and where I would be able to calculate, once and for all, how much money Haverford was spending to buy electronic versions of journals purchased in print. After I returned to my bitter reality, my colleague said "let's build it," and the idea for ERTS was born.

In addition to myself, the ERTS Team consisted of Swarthmore's Head of Technical Services, Bryn Mawr's Catalog and Serials Librarians, and the Tri-College Consortium's Special Projects Librarian. We decided early in the development process that the data we wanted to track fit into four categories: *licensors*, entities from whom we license e-resources; *vendors*, entities from whom we buy e-resources, *purchases*, expenditures made to access e-resources, and *titles*, individual e-resource titles. Because of our consortium status, ERTS needed the ability to both re-use data, and accommodate up to four instances of unique data, one for each of the libraries and one for the consortium. An example is the e-journal *Astrophysics*. A record for *Astrophysics* is entered in the titles file. Its licensor, Kluwer, is entered in the licensor file. Haverford buys the Kluwer collection through Palinet, in this case the vendor. Since each library has its own purchase arrangement with Palinet for the Kluwer collection to which *Astrophysics* is a part, each library creates its own purchase record, and is able to link to the shared title,

licensor, and vendor records. ERTS has really helped us get control of our e-resources in ways that just weren't possible before. It provides us with statistics not previously available, and the ability to keep in a central place all the data needed to manage our growing collection of licensed electronic resources. Additional information about ERTS, as well as its freely-available source code, is available at <<http://www.haverford.edu/library/erts/>>.

DLF ELECTRONIC RESOURCE MANAGEMENT INITIATIVE

There's a movement being sponsored by the Digital Library Federation that's trying to develop a solution to help everyone. The DLF Electronic Resources Management Initiative is a project to create standards for e-resource administrative metadata (Digital Library Federation, 2002). It is spearheaded by Tim Jewell, Head of Collection Management Services at the University of Washington, who in 2001 was conducting research for the DLF in this area. Tim talked with librarians who were building tools to help them manage e-resources, and concluded that this was an area where communication among developers would be beneficial. In November 2003, the ERMI Steering Group delivered a project report at the DLF Fall Forum in Albuquerque. In their report, they identified the goals of their work, the progress made towards these goals, and issues still in need of resolution. Some of these issues are the tension between open and proprietary systems, customization versus standardization, and the interoperability of stand-alone electronic resource management systems. Additional information about the project is available at <<http://www.diglib.org/standards/dlf-erm02.htm>>.

E-RESOURCE CONTROL

Providing free e-journal access to libraries that subscribed to the journal in print was a brilliant marketing move on the part of publishers, not unlike the drug dealer who generously gives away product until his target's hooked. Libraries turned their faculty and students into junkies, all while smiling from ear to ear, saying "aren't we wonderful." Libraries are in the midst of a time when they can no longer bankroll their users' habits, due as much to the uncertain economy as to changing pricing models.

Nearly half of Haverford's serials budget is spent on e-resources, yet my staff and I know very little about them. This concerns me a great deal. Too often users discover problems pertaining to e-resources before the library does. In the world of print serials, when an issue is late in publication, library systems notify us and generate a claim. This mechanism allows libraries to be proactive in determining the source of the problem, and to address it quickly, often before a patron is inconvenienced. On the other hand, when an e-resource is late in publication, or access has been unexpectedly and inappropriately turned off, it is typically a user who alerts the library of the problem. This scenario happens all too frequently, and puts the library in the dangerous position of being perceived as poor custodians of these expensive resources. I want to take back control of serials in my college. To this end, the Tri-College Consortium is investigating the feasibility of importing serial notification data as a way of getting a little more information than we have presently about our e-journal collections. We're also looking

at other aspects of our e-journal processes in an attempt to redesign them in ways that will eliminate redundancies and improve service to our users. How successful we'll be remains to be seen, but I know the time and effort we spend investigating these processes will be worthwhile.

CONCLUSION

Richard Atkinson, the recently retired president of the University of California, put it well when he said, "Librarians are now being forced to work with faculty members to choose more of the publications they can do *without*." (Atkinson, 2003). Atkinson's comments are in response to the ever-increasing subscription rates of scholarly journals. Weren't e-journals supposed to solve the serials pricing crisis? Unfortunately, the implication of electronic publishing has not been to lessen the cost of serials for libraries, but to increase them well above their already high and unsustainable rates of growth. Part of the dilemma is due to the perception that e-journals are complements, not replacements, to print journals. Although libraries make efforts to migrate subscriptions where feasible from print to electronic, all but a few exceptional institutions have hybrid, content-duplicative serials collections. In order to realize the transformation to electronic-only collections, we need certain publisher guarantees, and more problematic, the blessing of faculty, many of whom I've come to believe have only sentimental reasons, albeit passionate ones, for wanting to retain print journals. Life would be much simpler if publishers would cease their print operations entirely.

Significant changes are taking place that may finally free a portion of the scholarly literature important to teaching and learning. I refer to the open access movement, which is not a new concept. Various e-print servers in disciplines such as physics, computer science, and economics have existed for years. The Open Archives Initiative was founded in 1999 to enhance access to scholarship amassing in the growing number of e-print archives. Among the OAI's greatest accomplishments is its Protocol for Metadata Harvesting, a mechanism for providing and capturing metadata about e-prints. Theoretically, e-print repositories and a search/retrieval tool like the Protocol for Metadata Harvesting could obviate the need for traditional publishing outlets if peer-review controls were in place.

Last October, the Public Library of Science, a non-profit organization of scientists and physicians, established such controls when it launched its first journal, *PLoS Biology*. Three years in the making, the Public Library of Science is a fully-featured scholarly journal publisher, with an outstanding editorial team and a unique pricing model. Public Library of Science journals are free – almost. Readers of the journals pay nothing; authors pay a fee, \$1,500, to publish their articles – a fee the Public Library of Science expects to be subsidized by institutions and other agencies. To this end, the Wellcome Trust and the Howard Hughes Medical Institute recently announced that they would pay author fees for their researchers who published their grant-funded work in open access journals. That's a few hundred million dollars per year, research results of which may be available for free to libraries and the rest of the scholarly community. Not surprisingly, many commercial publishers are skeptical about the sustainability of an author-driven pricing model, reminding open access proponents that the PLoS was

awarded a \$9M grant from the Gordon and Betty Moore Foundation, and that \$1,500 per article, once the seed money is exhausted, is not enough to manage the costs associated with providing an electronic journal. One thing that seems clear, though, and which has been substantiated by the work of Steve Lawrence at the NEC Research Institute, is that open access scholarly journals, especially those with editorial boards the caliber of *PLoS Biology*, have a greater chance of being cited, and therefore of attaining a higher impact factor within their field (Lawrence, 2001). A high impact factor typically entices tenure-seeking faculty to publish within a journal, reversing the claim that freely-available web-based journals will never attract such scholarship. Obviously the open access movement is exciting, at least from a librarian's perspective, but there are still many issues yet to be resolved, both technically and politically. And we certainly have plenty of work to do in the meantime.

REFERENCES

- Atkinson, R. (2003). "A new world of scholarly communication." *Chronicle of Higher Education*, 50(11): B16.
- Cox, J. & Cox, L. (2003). *Scholarly publishing practice: the ALPSP report on academic journal publishers' policies and practices in online publishing*. Rookwood, UK: John Cox Associates.
- Digital Library Federation. (2002). "DLF Electronic Resource Management Initiative." Available: <http://www.diglib.org/standards/dlf-erm02.htm> (Accessed: 9 December 2003).
- Gorman, M. (1991). "The academic library in the year 2001: Dream or nightmare or something in between?" *Journal of Academic Librarianship*, 17(1): 4-9.
- Hitchcock, S., Carr, L., & Hall, W. (1996). "A survey of STM online journals 1990-95: the calm before the storm." Available: <http://journals.ecs.soton.ac.uk/survey/survey.html> (Accessed: 5 December 2003).
- Holmes, J. (1998). "Cataloging e-journals at the University of Texas at Austin: A brief overview." *Serials Librarian*, 34(1/2): 171-176.
- Lawrence, S. (2001). "Free online availability substantially increases a paper's impact." *Nature*, 411(6837): 521.
- McMillan, G. (1993). "Electronic journals: Access through libraries." In *The Virtual Library: Visions and Realities*. Westport, CT: Meckler.
- Moothart, T. (1996). "Providing access to e-journals through library home pages." *Serials Review*, 22(2): 71-77.
- Reynolds, R. (1995). "Inventory list or information gateway? The role of the catalog in the digital age." *Serials Review*, 21(4): 75-77.

Schulz, N. (2001), "E-journal databases: A long-term solution?" *Library Collections, Acquisitions, & Technical Services*, 25(4):449-459.

Spalding, C. S. (1957). "Keeping serials cataloging costs in check." *Library Resources & Technical Services*, 1(1): 13-20.