

## Cornell Law Library Scholarship@Cornell Law: A Digital Repository

---

Cornell Law Librarians' Publications

---

7-1-2005

# Rehabilitating Killer Serials: An Automated Strategy for Maintaining E-journal Metadata

David Banush  
*Brown University*

Martin Kurth

Jean M. Pajerek  
*Cornell Law Library*

Follow this and additional works at: <http://scholarship.law.cornell.edu/libpub>

 Part of the [Library and Information Science Commons](#)

---

### Recommended Citation

Banush, David; Kurth, Martin; and Pajerek, Jean M., "Rehabilitating Killer Serials: An Automated Strategy for Maintaining E-journal Metadata" (2005). *Cornell Law Librarians' Publications*. Paper 3.  
<http://scholarship.law.cornell.edu/libpub/3>

This Article is brought to you for free and open access by Scholarship@Cornell Law: A Digital Repository. It has been accepted for inclusion in Cornell Law Librarians' Publications by an authorized administrator of Scholarship@Cornell Law: A Digital Repository. For more information, please contact [jmp8@cornell.edu](mailto:jmp8@cornell.edu).

# Notes on Operations

## Rehabilitating Killer Serials An Automated Strategy for Maintaining E-journal Metadata

David Banush, Martin Kurth, and Jean Pajerek

*Cornell University Library (CUL) has developed a largely automated method for providing title-level catalog access to electronic journals made available through aggregator packages. CUL's technique for automated e-journal record creation and maintenance relies largely on the conversion of externally supplied metadata into streamlined, abbreviated-level MARC records. Unlike the Cooperative Online Serials Cataloging Program's recently implemented "aggregator-neutral" approach to e-journal cataloging, CUL's method involves the creation of a separate bibliographic record for each version of an e-journal title in order to facilitate automated record maintenance. An indexed local field indicates the aggregation to which each title belongs and enables machine manipulation of all the records associated with a specific aggregation. Information encoded in another locally defined field facilitates the identification of all of the library's e-journal titles and allows for the automatic generation of a Web-based title list of e-journals. CUL's approach to providing title-level catalog access to its e-journal aggregations involves a number of tradeoffs in which some elements of traditional bibliographic description (such as subject headings and linking fields) are sacrificed in the interest of timeliness and affordability. URLs and holdings information are updated on a regular basis by use of automated methods that save on staff costs.*

The enormous expansion of electronic journals available in full text, particularly those bundled in large aggregator packages, has been both a boon and a burden to libraries and users. The benefits of these resources for users—fast and nearly ubiquitous access to content, easy virtual browsing of individual issues and articles—are obvious. But the complexity of managing licenses, tracking expenditures, and providing accurate information about available titles and coverage has introduced challenges exponentially greater than the well-documented (and oft-lamented) variability of traditional print serials. Clearly, users want quick, convenient access to electronic journal content; certainly, libraries want to provide that access. The conundrum lies in identifying and maintaining the most effective means to inform users about what is available. A single

solution suitable for all contexts remains elusive.

Although frustrated by the difficulty of finding the magic bullet for simplified e-journal metadata management, librarians have not abandoned the search. Catalog librarians in particular have wrestled for some time with the problems surrounding e-journals. Catalogers know well the challenge of how best to provide access to these highly visible and popular, but extremely volatile, resources. In some institutions, traditional methods of serial bibliographic control—title-by-title cataloging, with each title viewed for verification of descriptive details, modes of access, and the recording of detailed holdings information—have continued, with varying degrees of success. In others, more streamlined processes have been adapted. In yet others, particularly smaller institutions with

David Banush (dnb8@cornell.edu) is Head, Bibliographic Control Services, Central Technical Services, and Martin Kurth (mk168@cornell.edu) is Head, Metadata Services Online Library, Cornell University, Ithaca, New York. Jean Pajerek (jmp8@cornell.edu) is Head of Technical Services and Information Management, Cornell Law Library.

few if any catalogers dedicated solely to serials, no catalog-level access is provided; staff are simply unable to maintain the catalog entries. National policies, particularly as outlined by the Cooperative Online Serials Cataloging Program (CONSER), have evolved as practices were tried, then altered or abandoned as changing circumstances and experiences dictated. Yet even as libraries have adapted practices to accommodate changing realities, certain assumptions about bibliographic record content (for example, the need to provide classification, standard subject headings, linking entries to related titles, and detailed descriptive notes) have remained largely intact.

Libraries, of course, have experimented with new approaches to bibliographic control for e-journals. Over the past decade, librarians have explored and implemented various alternative means to facilitate title-level access through the catalog to e-journals by using data that originate outside of the library. In the late 1990s, the Program for Cooperative Cataloging's (PCC) Standing Committee on Automation urged vendors to create MARC bibliographic records themselves and supply them to customers as part of the aggregator bundle. Some vendors have heeded the call, but many others have not, leaving significant gaps that libraries must cover with their own resources. Serials management companies like Serials Solutions, which emerged in response to librarians' (and end users') frustrations with licensed but uncataloged e-journals, have expanded their own services to include MARC records. They offer a way to fill in the coverage gaps created by those aggregators that do not supply bibliographic data along with their content. Yet these developments, while undeniably positive, come at a cost that many libraries cannot afford. They require some level of systems expertise, the ability to handle batch loading routines, and, of course, the financial resources not only to sup-

port the staff with such expertise, but also to pay any applicable subscription costs for the services.

Like other institutions dealing with intellectual access to electronic journals and the maintenance of titles and coverage data, Cornell University Library (CUL) has used various strategies, sometimes in accordance with national policies, and occasionally at odds with them, to provide the best possible service to users. While doing so, CUL technical services staff naturally have had to balance e-journal access with many other competing processing needs in times of static or declining staffing levels for technical services activities. This paper describes the work CUL staff have recently completed to develop and implement an e-journal management approach that assembles techniques developed locally and at other institutions into a comprehensive workflow for e-journal access and record maintenance. The strategy varies from CUL's own past practices, from current national practice as outlined by CONSER, and from traditional cataloging methods for serials titles. The method relies heavily on automation and brief bibliographic records. It challenges an implicit assumption of traditional serials bibliographic control by presuming that up-to-date title and coverage information is more important than full MARC cataloging, including access by classification and subject headings. Most importantly, the heavy use of automation and persistent identifiers in the records enables CUL staff to locate, extract, update, reuse, replace, or completely delete bibliographic data quickly and with relative simplicity in batches and obviates the need for one-by-one record processing by staff. The methods CUL has developed and employed in this strategy are scalable and applicable to other kinds of resources; indeed, they also may serve as a model for one method of automated metadata management for other institutions.

## **A Review of the Literature on E-journal Access**

Studying the library literature offers an opportunity to place CUL's current approach to e-journal access in an historical context. The history of library access to e-journals since the mid-1990s presents itself as concurrent movements from single-record to separate-record cataloging approaches, from manual to automated processes, and from fuller to briefer bibliographic records. Though libraries have sought to use the library catalog to provide more or less integrated access to print and electronic versions of journals during this period, the stand-alone e-journal database or Web list has continued to be a primary delivery mechanism for e-journal access. The single most influential factor in determining e-journal access strategies has been the emergence of the aggregator package as the dominant form for commercial distribution of e-journals. The mutability of aggregator holdings and the ease with which library selectors can add or cancel aggregator subscriptions have prompted library programmers to devise automated techniques that simplify the maintenance of e-journal records in sets. A close examination of the literature will help substantiate these claims.

Using the library catalog as a mechanism to enable users to discover and connect to e-journals dates from about 1994. Around that time, early Web-based catalog interfaces began to appear, providing direct links to e-journal content via the 856 electronic location and access field that MARC developers had added to the MARC bibliographic format in 1993.<sup>1</sup> As end user and library staff interest in e-journals increased in the 1990s, many library practitioners responded by reasserting their belief that the library catalog should serve as the central site for access to all library resources, regardless of format.<sup>2</sup> At the same time, other practitioners favored Web lists

or databases that segregated and highlighted e-journals for easy access, thus providing separate access mechanisms for e-journals that were successors to the printed lists of serial title and holdings information that libraries had historically provided as complements to catalog entries for serial titles.<sup>3</sup> The experience of the Perry Library at Old Dominion University clearly exemplifies such an evolution from printed periodical list to Web list to online periodicals database.<sup>4</sup> By 1998, catalog and noncatalog access mechanisms had become commonplace. At that time, Shemberg and Grossman noted that 78.8 percent of institutions belonging to the Association of Research Libraries (ARL) and 39 percent of non-ARL libraries surveyed offered online catalog access to e-journals; at the same time, 87.1 percent of ARL libraries surveyed and 48.8 percent of non-ARL libraries surveyed used Web lists for e-journal access.<sup>5</sup>

In those libraries choosing to add e-journal holdings to their catalogs, library technical services operations accepted the responsibility of creating programmatic processing workflows to increase catalog access to e-journals. Early timesaving strategies often employed manual cataloging processes, such as those at Auburn University and the University of Texas at Austin, to add 856 fields and other fields related to electronic access to libraries' records for their e-journals' print counterparts.<sup>6</sup> The University of Pennsylvania's decision to perform manual separate record cataloging for electronic versions of print journals ran counter to the more common manual single-record workflows.<sup>7</sup> Following the explosion of e-journal access in aggregations, however, all manual e-journal cataloging approaches became problematic. Authors such as Calloun and Kara in 1999, followed by Jones in 2001, expressed their preference for automated ingest and maintenance of e-journal catalog records as expedient strategies for technical services units

seeking to keep pace with the acquisition and de-acquisition of e-journals in unprecedented numbers.<sup>8</sup>

The availability of e-journal metadata from external sources and the desirability of using automated methods to load and maintain record sets for aggregated e-journals have led many libraries, exemplified by the University of Tennessee, Knoxville, and the University of Glamorgan, to prefer adding separate records for e-journal titles, separate, that is, from the records for those titles' print equivalents.<sup>9</sup> In 1999, Martin and Hoffman surveyed forty-three Research I and Research II academic libraries to study how they provided access to e-journals from aggregators.<sup>10</sup> Of the libraries that added catalog records for aggregated e-journal titles, 20 percent used the single-record approach (combining e-journal information and print-journal information in a single-record), 16 percent added separate records, 9 percent used both methods, and 30 percent gave no indication of the approaches used.<sup>11</sup> The PCC Standing Committee on Automation Task Group on Journals in Aggregator Databases in early 2000 further legitimized automated, separate-record approaches to e-journal cataloging by recommending automated methods for creating e-journal record sets derived either from MARC print-journal records or from non-MARC e-journal metadata supplied by e-journal or third-party vendors.<sup>12</sup> The task group's report was also significant because it offered strategies for creating less-than-full-level MARC records for e-journal titles. That library reliance on automated handling of vendor-supplied e-journal metadata had become a growing trend was reflected in an informal Research Libraries Group (RLG) survey from January 2003.<sup>13</sup> The survey revealed that nine of twenty-three RLG member respondents used external sources for e-journal record sets, and that seven of the fourteen member libraries that did not use external sources

were either planning to do so or were investigating their use.

Though more and more libraries began using vendor-supplied metadata in automated, separate-record cataloging methodologies for e-journals, libraries did not universally adopt such approaches. Stalberg reports, for example, that library staff at St. Joseph's University elected to do manual, single-record cataloging for their aggregated e-journals because records for them were not available from external sources; further, St. Joseph's staff combined the information for e-journals offered by more than one aggregator into a single record because they believed that users preferred to retrieve one record per e-title.<sup>14</sup> Other libraries, such as Hong Kong Baptist University, implemented multiple workflows involving a mixture of manual and automated processes.<sup>15</sup> Hong Kong Baptist staff performed manual, single-record cataloging for print journals and e-journals in aggregations whose holdings tended to be stable, while they used an automated workflow for e-journals in unstable aggregations that collocated all versions of an e-journal in one bibliographic record. The cataloging community came to call the latter approach (that is, using manual or automated processes to create a single bibliographic record for all electronic versions of a given title) "aggregator-neutral" cataloging when CONSER adopted it as policy in 2003. The next section of this paper presents a detailed discussion of CONSER policies on e-journal cataloging.

Libraries using automated approaches to e-journal cataloging varied in the fullness of the MARC records they loaded into their catalogs. Using a mixed-level approach, Hong Kong Baptist University added some brief e-journal records to their catalog, but they loaded vendor-supplied full MARC records when they were available.<sup>16</sup> In a single-level approach, the University of Tennessee, Knoxville, and the Western North

Carolina Library Network (WNCLN) added brief records exclusively.<sup>17</sup> The records were created from their vendor-supplied non-MARC source data. In an alternative approach to choosing between full and brief records, WNCLN staff looked for ways to augment their brief records with classification data and subject terms derived from classification numbers.<sup>18</sup>

Though many libraries devoted a great deal of effort to adding e-journal holdings to their catalogs, they also continued to use non-catalog methodologies, such as Web lists and standalone databases, to facilitate e-journal access.<sup>19</sup> Auburn and Hong Kong Baptist Universities offered e-journal access via the library catalog, yet they also extracted catalog data to create Web lists of their e-journals.<sup>20</sup> Librarians at Yale University created “jake” as a cooperative database that collocated access data for e-journal full-text and indexing sources.<sup>21</sup> In what may have signaled an emerging trend in e-journal access, the Colorado Alliance of Research Libraries expanded on the jake database to develop an architecture that supported a hybrid approach, relying on both the library catalog and an external database of e-journal information. According to Meagher and Brown, developers of the Colorado Alliance’s Gold Rush database envisioned that libraries using Gold Rush would link from 856 fields in MARC e-journal records to a Gold Rush display containing links to aggregators that offer full text for a given title as well as to abstracting and indexing services that index it.<sup>22</sup>

Most practitioners writing about e-journal access methodologies have addressed their libraries’ attempts to develop e-journal processing strategies that sought to minimize the effort needed to maintain e-journal records over time. That the need to update e-journal records is a legitimate concern for libraries is reflected in the Western North Carolina report that more than one-third of the 8,000 e-journal

records supplied by their vendor contained changes in the first bimonthly update file they received.<sup>23</sup> Libraries have devised a variety of techniques to respond to such volatility in e-journal access. In the manual approach used at St. Joseph’s University, staff added separate note fields for information corresponding to different e-journal versions in order to simplify record maintenance.<sup>24</sup> Hong Kong Baptist University staff placed unique identifiers for e-journal titles in 035 fields to enable monthly overlays of e-journal records.<sup>25</sup> Western North Carolina programmers wrote record matching scripts that allowed record overlays and field deletions.<sup>26</sup> And, finally, Oregon State University developers designed a local application that uses MARC 001 field matching to delete vendor-supplied records for dropped titles.<sup>27</sup>

A general trend emerges from the overview of the literature on e-journal management—namely, the use of externally supplied metadata in automated processes to create MARC records in varying levels of richness that can be modified or removed in record sets. The e-journal management innovations and variations reported in the library literature reflect how important straightforward e-journal access is to libraries and how challenging it is for libraries to provide that access.

### **E-journal Cataloging Practices CONSER Policy**

Even with the trend toward more automated processing and externally supplied e-journal descriptive information, traditional cataloging has continued to play a significant role in many libraries’ e-journal management strategies. But traditional bibliographic control practices for serials have not remained static. Indeed, a look at the evolution of national-level cataloging practices as outlined by

CONSER reflects the challenges that e-journal description and access have posed to efficient, effective bibliographic control. Over the past decade, CONSER has endorsed various approaches for contributing e-journal cataloging records to its database, most recently (since July 2003) the aggregator-neutral record. The shifts in policy have naturally been influenced by changes in the MARC 21 standard, AACR2, the limits of public displays offered by integrated library systems vendors, and, of course, catalogers’ increasing levels of experience with e-journal cataloging.

Prior to the July 2003 implementation of the aggregator-neutral record, CONSER guidelines offered libraries the option of creating a separate bibliographic record for an e-journal that also exists in print format, or of combining information about print and e-versions in a single bibliographic record. Under these guidelines, a print serial title and multiple electronic versions of the same title would appear on a single record if that record also covered the print version. However, if a cataloger chose to catalog print and electronic versions separately, the guidelines required that a cataloger create a separate record for each electronic version of a serial issued by a different distributor or aggregator. Previously, CONSER guidelines did not permit a record describing an electronic serial with reference to multiple aggregators that did not also contain the description for the original print version.

As e-journals proliferated, and an increasing number of titles became available through more than one aggregator, CONSER decided that its policy of creating a separate record for each aggregator’s version of an e-journal made these records “confusing and hard to maintain.”<sup>28</sup> That the creation of separate records for different electronic versions of the same title “increase[s] the likelihood of inadvertent duplication, frustration for

searchers, and irritation and confusion for all concerned” was becoming increasingly obvious.<sup>29</sup> CONSER decided, therefore, that the time had come to rethink the separate record policy with an eye toward providing less-confusing catalog records and minimizing the need for local editing.

CONSER’s most recent solution to the problem of multiple records for different versions of e-journals is the aggregator-neutral record. The aggregator-neutral record is a bibliographic record that is “separate from the print [and] that covers all versions of the same online serial on one record.”<sup>30</sup> Under the current policy, catalogers would represent the electronic versions of a title, such as *ABA Banking Journal*, which is available from at least four different aggregators, with a single catalog record instead of four separate records. Because they reflect the title at a more abstract work level, aggregator-neutral records lack certain elements of description associated with separate records for specific iterations of e-journals, such as a uniform titles qualified by aggregator name. The goal is to present the searcher with one-stop shopping for all electronic holdings of any given title for which it has a license to provide access. At the time of this writing, OCLC is presently using a combination of automated and manual processes to collapse records in the CONSER database and edit them to conform to the guidelines for aggregator-neutral records.

### Cornell University Library Practice

The rapidly changing nature of e-journal publishing and the evolution of cataloging standards intended to accommodate the changes spearheaded by CONSER have led to the development of practices at CUL that diverge from CONSER’s recently implemented policy. In the mid-1990s, the “single record versus separate records” question was debated at length at Cornell, as it was at many other libraries. CUL

developed local guidelines that allowed for the creation of combined print and electronic records in cases where CUL’s holdings included the print version of a title although the local policy stated a preference for separate records unless there was a compelling reason not to create them. In the early days of e-journal cataloging, catalogers at CUL had the luxury of being able to conduct detailed analyses of e-journal aggregators to determine the most efficient and cost-effective way of cataloging them. Decisions on whether to create separate records for e-versions or to use the combined record approach were based on factors such as the size of the collection, the percentage of the collection owned by CUL in print form, the availability and completeness of bibliographic records, the amount of local editing required, and the feasibility of batch-loading the records.

When an e-journal was supplied by more than one source, early CUL policy was to create a single holdings record with an online location representing all the e-versions. The holdings statement was compressed to reflect the combined coverage offered by the multiple providers. For example, the *American Journal of Philology* is part of both JSTOR (for back issues) and Project Muse (for current issues). The holdings statement for the e-version of the title conflated the coverage into a simple statement, v.1 (1880)- . The resulting OPAC display allowed users to see at a glance that the library’s electronic holdings went all the way back to the first volume of the publication. Eventually, CUL staff discovered that the lack of granularity in these combined holdings statements presented maintenance problems when, for example, one of the providers discontinued or changed its coverage and a cataloger had to determine where that provider’s coverage ended and another’s began.

Being able to spend the time to do so, however, was a rare occurrence; CUL did not generally undertake sys-

tematic maintenance of e-journal cataloging records. One exception was the ProQuest database, one of Cornell’s largest and most heavily used aggregations. High-level staff maintained the ProQuest data manually. In addition, staff outside of Central Technical Services did manual maintenance on a few small or medium-sized aggregations, using updated information supplied by vendors. CUL technical services staff updated other e-journal records on an ad hoc basis, typically in response to reports of problems from public services staff or library users.

By spring 2001, e-journal publishing was expanding at an exponential rate and was undergoing changes that library technical services staff could not afford to ignore, such as the practice of some e-journal suppliers to limit access by imposing embargoes or moving walls on their coverage. Catalogers no longer had time to create and maintain individual records manually for each title in every aggregation purchased by the library, or even to update the existing print records with additional e-journal information. E-journal publishing had outstripped the library’s ability to keep up with it using traditional cataloging methods. CUL technical services management felt that the volume of e-journal cataloging and maintenance called for a new approach to e-journal bibliographic control. The technical services managers decided to provide title access to large numbers of e-journals in aggregations by creating and adding to the catalog abbreviated, machine-generated bibliographic records, dubbed “sleek” records. At the time, the library anticipated that full-level cataloging would eventually be supplied to replace the sleek records. However, resources have yet to become available to upgrade the sleek records to full-level records.

Initially, technical services staff used a locally developed program to generate sleek records from title lists supplied by vendors. In fall 2001, CUL contracted with Serials Solutions

to purchase title-level bibliographic data for e-journal aggregations not yet cataloged and for updated data for those e-journal sets already cataloged. Every two months, CUL received a spreadsheet from Serials Solutions with the journal title, International Standard Serials Number (ISSN), start date, end date, provider, and URL. Information technology staff in Central Technical Services converted the Serials Solutions-supplied spreadsheet into a tab-delimited text file and ran the data through a locally developed Perl (a high-level programming language) script to generate pseudo-MARC records. These records were then converted into MARC by using a utility called MARCEdit.<sup>31</sup> The resulting records were then loaded into the CUL catalog by means of a customized Visual Basic program.

CUL technical services staff assigned the sleek records a MARC encoding level 3 ("abbreviated level") and elected not to export the records to the bibliographic utilities. The records included the title information, URL, coverage information, and aggregator information extracted from the Serials Solutions spreadsheet. They did not include subject headings, call numbers (or classification numbers), information pertaining to preceding or succeeding titles (780, 785 fields), or the availability of other formats (530, 776 fields).

Occasionally, because of interface or content changes in a given aggregator package, an entire set of sleek records had to be removed, as when the Dow-Jones Interactive package became Factiva. This process was facilitated by including a special, searchable 899 field in the sleek records that identified the aggregator associated with the title. Staff populated the 899 field using a controlled vocabulary of abbreviations or codes, one specific to each aggregator or provider. The 899 was added manually or via automation, depending on how the record was created. The 899 field is illustrated in figure 1.

The machine-generated sleek records solved some of CUL's e-journal cataloging problems. The library was able to provide title-level access to its e-journals through the catalog, but record maintenance was still an unresolved issue, as CUL sought a method for automated maintenance as well as automated record creation. The CUL database still included e-journal records that had been created at different times, reflecting a variety of sometimes contradictory rules and policies. Applying an automated, across-the-board maintenance routine to this disparate set of records would prove to be challenging. In summer 2002, library management formed a committee to address the increasingly complex issues associated with e-journal cataloging, particularly the need for a systematic approach to ongoing maintenance of e-journal catalog records.

### The CUL E-journal Maintenance Task Force: Goals and Objectives

In July 2002, the CUL Technical Services Executive Group (TSEG)

created the Electronic Journal Maintenance Task Force to examine the library's policies on e-journal access and to recommend new strategies for maintaining the collection of electronic journals to which the library provides access. As noted above, CUL technical services units did not necessarily coordinate their handling of e-journal maintenance. In fact, staff in the several processing centers took various approaches to maintenance, from manual efforts to the use of Serials Solutions data. Although an in-house manual covered cataloging issues for electronic resources generally, CUL had no true institution-wide maintenance policies or best practice in place.

This bric-a-brac approach reflected the generally decentralized processing environment at CUL. For TSEG, which represents all of the CUL processing centers and sets technical services policy at the system level, an uncoordinated, scattershot maintenance strategy was no longer desirable. The group wanted a more cohesive approach. TSEG wished to ensure better service to users while simultaneously rationalizing

---

022 __	a 0020-7047	
035 __	a 4127056	
040 __	a NIC  c NIC	
245 00	a International Journal for Philosophy of Religion  h [electronic resource]	
260 __	a [S.l.] :  b Kluwer Journals Online	
362 0_	a 1997-	
500 __	a Title from Serials Solutions journal URL list.	
538 __	a System requirements: Internet connectivity, World Wide Web browser and Adobe Acrobat reader.	
538 __	a Mode of access: World Wide Web.	
856 40	3 Current issues  u <a href="http://encompass.library.cornell.edu/cgi-bin/checkIP.cgi?access=gateway_standard%26url=http://www.wkap.nl/jrnltoC.htm/0020-7047">http://encompass.library.cornell.edu/cgi-bin/checkIP.cgi?access=gateway_standard%26url=http://www.wkap.nl/jrnltoC.htm/0020-7047</a>	
899 _0	a kluwerj	←
906 __	a gs	
946 __	a 1997-	
948 0_	a 20011204  b i  d sleek  e cts	
948 1_	a 20011204  b s  d sleek  e cts  f ?	
948 2_	a 20030319  b m  d batch  e cts  f j	

Locally-defined 899 code indicates the aggregator to which the title belongs.

---

Figure 1. A sleek record in the CUL catalog.

the effort expended in handling and managing maintenance systemwide. The charge that the group drafted for the task force called for the work to take place over the course of one year. Among other things, the charge instructed the four-member task force to examine past and current cataloging practices for e-journals; explore their maintenance implications; assess the feasibility of using vendor-supplied MARC records; develop a plan for creating a title list of all e-journals that could be derived from the library catalog; and, most importantly, create a plan for the ongoing maintenance of e-journal bibliographic and holdings data in the Cornell catalog.

A unifying, though implicit, theme running through the various specific actions the charge put forth was clear to the task force: achieve the best possible result with maximum flexibility and the most parsimonious use of financial and human resources. The underlying assumption of the charge was that existing maintenance efforts were both too expensive and too limited to be justifiable or sustainable. Users, public services staff, and technical services managers seemed to agree, if only tacitly, that the volatile nature of e-journals required a vastly faster and more efficient approach than any manual maintenance effort could provide. Moreover, task force members concluded that, given a very tight financial situation in the library generally, receiving additional funds for purchasing records from outside sources or for hiring additional staff to handle manual maintenance were not realistic scenarios. Thus the task force quickly determined that the solution most likely to win the backing of TSEG would be heavily automated and would make use of tools and data already available. The group carried out all subsequent analyses and formulated the potential strategies in that spirit.

### Identifying E-journals in the CUL Database

Several members of the task force worked to identify the complete set of e-journals within the CUL catalog. This ostensibly simple task proved more daunting than originally anticipated, in part because of the variations in both local and national cataloging practice over the years and also because CUL had never used any locally defined code as a marker or identifier for e-journals. Selecting records for e-journals from the CUL database was only possible by using standard data in the MARC bibliographic and holdings records. However, variations in practices made accounting for all of the possible combinations of MARC fields in CUL bibliographic and holdings records complicated. Though the group had members with considerable expertise in the use of Structured Query Language (SQL), a standard computer language for accessing and manipulating databases, devising the most effective and comprehensive means to identify the body of e-journals was a challenging undertaking. After numerous attempts, the group concluded that a concatenated series of five Microsoft Access queries against the CUL Voyager database successfully identified the subset of e-journal records.<sup>32</sup>

Once the queries had identified the records, the task force then proposed a method to code them. Unique coding would facilitate their identification and extraction in a more rapid, flexible manner than the SQL queries could provide. After considerable discussion, the group decided to use a locally defined bibliographic record code to specify e-journal titles. The principal advantage of the approach was that the use of locally defined MARC fields in the bibliographic record had become standard practice at CUL. Moreover, CUL staff had no comparable means of harvesting

such data from other sources (such as holdings records), and creating the mechanisms to do so would have entailed higher opportunity costs than the task force was willing to assume. Since 2000, CUL staff had been using the locally defined 948 MARC bibliographic field for statistics-gathering purposes, and a subfield (f) and value (e) had previously been defined in that field for electronic resources generally. The task force defined a new value for e-journals (j) to distinguish them from other e-resources and proposed to implement the new coding scheme in two phases: first prospectively (and manually), as catalogers added new titles to the database, and then retrospectively through automation, adding the codes to the bibliographic records previously identified via the queries. This work was completed by March 2003.

The ability to extract e-journal bibliographic data from the catalog quickly allowed the task force to meet another of its objectives: generating a Web-based title list on demand of all e-journals available from Cornell. Data extracted from the MARC records could now be used for creating such a list. Previously, the library relied on data from Serials Solutions to create its title list, but the investigations done in identifying e-journal records in the catalog indicated that more than 4,000 titles, or nearly 25 percent of all e-journals, were not covered by Serials Solutions and were thus unrepresented in the list. The majority of these titles were items that were free or available from small institutes, government agencies, professional societies, or independent publishers of varying types. Although another group of staff developed the specific plans for generating the list and displaying it, the task force's work laid the foundation for the inclusion of these resources in the title list alongside their more commercially prominent counterparts.



### Data Analysis, Cleanup, and Preparation

With the coding to identify e-journal records in place, the task force began a series of analyses to study those records in more detail. Among the most important questions the group needed to answer were how many e-journal titles had been cataloged as separate records and how many had been added to their print counterparts on a single record. CUL cataloging staff had employed both practices, depending on the prevailing national and local consensus at the time, the availability of staff resources, or relative institutional priorities. Over the course of time, a separate record policy had emerged as a default as a result of the sleek record approach, but staff had handled many large aggregator sets (such as ProQuest) using a single-record method. Moreover, some single records represented both print (or print and microform) and multiple aggregators' versions of the resource. Exactly how many titles had been done one way or another was unknown. The task force was interested in standardizing the database so that when the library implemented its new workflow, the data would reflect as much internal consistency as possible. The group understood that without a consistent policy of using either single or separate records, the application of automated solutions for e-journal maintenance would become much more complicated.

Coding the full set of CUL e-journals with identifying values made extracting and analyzing them simple. To determine the number of titles given single-record treatment (internally referred to as multiple version, or "mulver," records), the task force first had to determine which bibliographic records for e-journal titles also had holdings records indicating that both print and electronic holdings were attached. Bibliographic or holdings records that were suppressed from

public view were considered inactive and were therefore ignored in the data harvest and analysis. Using those rough criteria, the task force was able to identify nearly 3,500 titles that had been done as single (mulver) records. A small subset of these (fewer than 400 titles) represented "multi-mulvers," that is, single records with more than one link to electronic text. Separating these records manually—moving to a separate record for the electronic version—would have been a significant task. The group estimated that cleaning up a single record ("demulverization") and creating a new separate record for the electronic version, moving holdings, and other clean-up tasks would take an experienced staff member roughly twenty minutes. Thus, the group calculated that a manual clean-up of all 3,500 mulver titles would have involved more than 1,100 staff hours, or 27.5 work weeks for one full-time staff member.

Task force members felt that this amount of time was unacceptably high. Instead, the group chose to explore an automated approach to record clean-up, with manual efforts reserved only for those titles that would fall outside automation's reach. After consulting with appropriate staff about the feasibility of returning mulver records to their print-only state programmatically, group members develop a detailed sheet of specifications for such a routine. The specifications called for removing certain fields from the bibliographic record that pertained to the electronic version (such as 538 mode of access notes, 506 restrictions notes), as well as the holdings record associated with the electronic version of the item. The coding identifying the title as an e-journal record was also removed. Testing of the routines in the CUL test catalog proved encouraging, and the group moved forward with a batch job that quickly cleaned up the mulver records.

To address the titles that fell outside of the automated cleanup routines,

the task force wrote various queries against the set of all e-journal records. The reports from these queries contained the titles, record ID numbers, and other data. Task force members forwarded the information on to cataloging staff for manual cleanup. Several hundred records were handled in this fashion. Because the number of titles requiring manual cleanup represented a very small percentage of the overall number of titles, CUL cataloging staff completed most of this work in a period of several weeks.

### Current Approach to Cataloging and Maintenance

The maintenance and cataloging policy laid out by the task force centered on separate record cataloging for each title, including individual records for each electronic manifestation of a title. The approach essentially extended the existing sleek record approach to all titles for which external metadata was available. The task force had many spirited discussions, not only about the best way to approach the issue of separate records and coverage from multiple aggregators, but also about other consequences of the decision to use automated e-journal record maintenance. Ultimately, the group elected to use a separate record approach, with one record for each version or expression of the title. Thus, for titles provided by multiple aggregators (such as JSTOR, ProQuest), CUL represents each version with a separate record and holdings statement. The task force decided to follow this method because the use of completely separate records would simplify automated maintenance routines. The separate records, with their e-journal and 899 codes in the locally defined MARC fields, also make identifying all of the records provided via a given set a relatively simple task, should that provider be

dropped or have a blanket change in coverage.

The maintenance process consists of two separate but interdependent steps. First, using title and holdings data provided by Serials Solutions, the information technology librarian generates a series of brief catalog records for each title. Many of the e-titles have print version records in the CUL catalog. Moreover, in certain instances, the library also receives the same title (though not necessarily the same coverage) from multiple aggregators. To assist users in making sense of the resulting OPAC displays, the task force recommended generating a uniform title for each record created from the Serials Solutions data. A series of conditional statements was built into the routines for generating the MARC data, adding or editing (as appropriate) a 130 field using the title proper and a parenthetical qualifier for the aggregator. The group felt the use of these titles, though not strictly in accordance with current cataloging codes and practices, would help users distinguish among the different versions of the titles available in the catalog. Figure 2 illustrates a CUL sleek record that contains a machine-generated 130 field.

The automated routine adds note fields, but only those that apply across all titles, such as access restrictions, basic system requirement notes, and source of title notes. Holdings records using the coverage dates and at the level of specificity provided by Serials Solutions are also created at the same time. In addition, for licensed resources, the 856 field is modified to include a prefix that indicates one of two different authentication types—all of Cornell, or all of Cornell except the Weill Medical School campus in New York City. A code that represents the authentication type is stored in the locally defined 906 field of the MARC record.

That these records are brief deserves emphasis. Given the limitations of the source data, some of the common fields found in standard serial MARC records are omitted. For example, the routines cannot assign classification numbers, even at a general level. The program cannot supply title linking fields (77X, 780/785) or any title-specific notes. For aggregators whose coverage is known to be limited to a particular range of dates in the publication's history, a generic note is generated for public display. In the case of JSTOR, for example, with its moving wall of con-

tent coverage, the program adds the note "Most recent issues not available. Please check resource for coverage." The message alerts catalog users that coverage restrictions apply, but does not specify if the moving wall is for three or five years. Because the text appears as the hyperlink to the content in the CUL Voyager catalog, users can readily see it in the results display. All of the brief MARC records are given encoding level 3 (abbreviated level), and they are not exported to the bibliographic utilities.

Although used extensively, CUL does not rely on Serials Solutions e-journal metadata for every title. For ProQuest, CUL loads the free, full MARC data set provided by the vendor. ProQuest staff adapt these records from existing records representing their print counterparts. CUL staff take that file and perform a preloading routine to add particular notes (such as 506 restrictions notes) or to remove unwanted fields before loading them into the catalog. As with the Serials Solutions-based records, holdings data is loaded as reported by the vendor. Using similar criteria, the same routine used for the Serials Solutions titles also creates uniform titles for ProQuest records. Linking fields are often present in these records, but may not be in each record; CUL staff do not check for the presence of the data or verify its accuracy if it is supplied. Because many ProQuest titles have embargoes on coverage, the task force specified that the program should add a generic note ("Most recent issues may not be available. Check resource for coverage.") to each record. The generic disclaimer makes no attempt to determine if an embargo applies to any particular title. A further step is replacement of the 856 value from ProQuest with a locally created persistent URL (PURL). The ENCompass system, which CUL uses as its platform for the e-journal title list, has a character limit of 255 characters.

---

```
022 __ |a 0569-4345
040 __ |a NIC |c NIC
130 0_ |a American Economist (New York, N.Y. 1960 : Online : Business Source Premier)
245 10 |a American Economist |h [electronic resource].
260 __ |a [S.l. : |b s.n.]
362 0_ |a 1964-
500 __ |a Title from Serials Solutions list.
506 __ |a Access restricted to licensed institutions.
538 __ |a System requirements: Internet connectivity and World Wide Web browser.
538 __ |a Mode of access: World Wide Web.
856 40 |u http://encompass.library.cornell.edu/cgi-bin/checkIP.cgi?access=gateway_standard%26url=http://search.epnet.com/direct.asp?db=buh&jid=%22AEC%22&scope=site |z Connect to full text.
899 __ |a BusSourcePrem
906 __ |a gs
948 0_ |a 20031211 |b i |d batch |e cts
948 1_ |a 20031211 |b s |d batch |e cts |f j
```

---

**Figure 2.** CUL sleek record with machine-generated 130 field added.

Because many ProQuest URLs exceed that limit, this step is necessary to load the records into the ENCompass e-journal repository.

Titles covered by neither Serials Solutions nor ProQuest are given full catalog records and are handled manually. CUL selectors complete a networked electronic resource selection form to initiate the cataloging of new titles. Acquisitions staff winnow out titles available from Serials Solutions or ProQuest and pass the remaining selections on to cataloging staff for handling. Acquisitions staff also search all such titles in the bibliographic utilities for cataloging copy; if found, the copy is edited as appropriate for inclusion in the CUL catalog. Resources lacking copy are given full, original records. These records include the local MARC coding identifying them as e-journals, but in most cases lack the 899 field codes that associate them with a particular aggregator set. Descriptions and holdings are based on viewing the resource itself and are created according to national practices. The manually created records are given the appropriate encoding level and are exported to the bibliographic utilities along with most other newly cataloged CUL resources.

CUL has now largely automated maintenance for e-journals. Several times a year, updated ProQuest records (obtained from the vendor) and a refreshed data set from Serials Solutions are compared with the records in the CUL catalog that were generated from those sources via machine matching. This automated mechanism provides regular updates to more than 80 percent of the more than 25,000 titles in the e-journal collection.

The process runs in a series of steps. Title and coverage data are maintained within the Serials Solutions Web interface. The Serials Solutions database is the database of record for most of CUL's licensed electronic journal sets. Approximately

every two months, Serials Solutions sends CUL an updated file reflecting changes made to titles and coverage information. When the file is received at Cornell, it is converted into a tracking table within a Microsoft Access database. The table records the information supplied, as well as other administrative metadata, including whether the particular aggregator set is included in the automated workflow. (Some sets, such as HeinOnline, have data available but are not refreshed, as noted above.)

For titles in aggregations that CUL continues to license, MARC records are generated from the Serials Solutions data set. The resulting file is placed on a local file server. Then, a process is run that compares the new records within each aggregator set (based on the aggregator code in the 899 field) with existing records in that same set from the Voyager catalog. Based on a title match within the set, the 856 and coverage data are compared. If they are different, the record is updated with the new data. If a record is in the new set, but not already in the catalog, then the new record is added. If a record already exists in the catalog, but is not in the new Serials Solutions file, then the record in the catalog is marked for deletion.

When a new aggregator package is added to the CUL collection, the process enables the rapid creation of records for that set. New records sets are processed on demand. This involves running a custom Perl script that converts the Serials Solutions data into MARC records, then loads them into the Voyager catalog through calls to Endeavor's BatchCat API (Application Program Interface).

Appropriate 899 aggregator codes are added to the records before loading; they are then placed in the routine automated maintenance queue.

For titles that fall outside of the automated routines, only passive maintenance is performed. That is,

after a title has been added to the catalog, no regular checking is done, and records are updated or deleted only after reports of problems from users or public services staff. At the time of this writing, CUL does not employ a URL checker, though the implementation of such software has been discussed.

### E-journal Title List

The development of an automation-rich maintenance routine was a significant milestone in the E-journal Maintenance Task Force's work, but it was not the end of that work. TSEG had also charged the team to create a Web list of e-journals based on data in the CUL catalog. Although CUL had maintained a Web list of online e-journal titles via the Cornell University Library Gateway from the late 1990s onward, the list had only contained data provided by Serials Solutions. The older Web list did not, therefore, reflect more than 5,000 CUL e-journal titles not covered by Serials Solutions but indexed in the catalog. Thus, users who relied on the Web list rather than the catalog were unaware of a very substantial number of both licensed and free-access journals in electronic format.

Task force members agreed that a complete, accurate Web list of e-journals would represent a significant step forward in e-journal access for CUL end users. The 899 codes to identify e-journals and aggregators added as part of the new maintenance routine also were intended to permit library staff to extract data from MARC e-journal records for use in other applications, either for public access or administrative support. Once CUL technical services units implemented the e-journal and aggregator coding, library staff were able to extract some or all of the catalog's e-journal records with straightforward databases queries. Information tech-

nology staff in technical services used the harvested MARC data to generate brief e-journal entries for a new Web list called "Find E-journals," which offers searching and browsing functionality via Endeavor's ENCompass digital library management system. In addition to using the ENCompass system for the Find E-journals Web list, CUL uses ENCompass for its "Find Databases" service, which provides access to approximately 1,000 electronic reference sources, and for "Find Articles," which enables federated searching of selected abstracting and indexing databases.

The CUL technology specialists created a series of scripts to draw from MARC fields 110, 130, 240, and 245 to construct individual title headings for the Web list that generally follow the syntax of uniform titles. Additional scripts convert e-journal metadata from MARC to a local implementation of Dublin Core that contains title, identifier (URL), relation (aggregator code), and bibliographic record number elements. The title element in Find E-journals also contains coverage information appended to the title portion of the element content. The title is the only element that displays to the public. An example of one of the XML-encoded records for Find E-journals appears in figure 3.

To ensure harmony between Find E-journals and the catalog, CUL technical services staff have established a workflow that updates the catalog, extracts catalog data, encodes catalog data in XML using the Find E-journals element set, and loads Find E-

journals records into ENCompass for delivery to CUL users. Putting this workflow in place relieved the uncertainty that staff and users had had about whether the catalog or the Web list offered more complete access to CUL e-journals. It also has simplified the answer that technical services staff can give when asked about the differences between the two e-journal access methods: the e-journal content in both systems is now the same.

### Drawbacks and Benefits of the CUL Approach

CUL spends hundreds of thousands of dollars to purchase access to e-journal aggregations. Such an investment would not be justified if library users were unable to access these resources easily and conveniently. In considering how best to facilitate e-journal use, the library found itself confronted with the need to reconcile incompatible priorities: to create and maintain bibliographic records for expanding sets of e-journals while also being asked to cut or reallocate significant portions of its operating budget. Technical services managers realized that doing both concurrently could only be accomplished through automated means, and that the approach could force CUL to make some difficult compromises. In the course of determining the best possible automated solution for the creation and maintenance of e-journal records in the CUL catalog, the E-journal Maintenance Task Force considered a number of possible strategies,

each entailing its own set of advantages and disadvantages. Clearly, no single strategy would be entirely satisfactory to all of the library's constituents. But the need to provide title-level access through the online catalog using an automated procedure with minimal human intervention was one of the group's guiding principles. The solution the task force arrived at involves a number of tradeoffs, but critical library stakeholders believe that the benefits to users outweigh the disadvantages.

What are the negatives and positives of the CUL e-journal maintenance approach? As with most applications of bibliographic control, the pluses and minuses of the e-journal workflow are relative to the various constraints and resources in the CUL context. What one considers a drawback or a benefit is a function of perspective. That is, others may disagree as to both the kind and degree of seriousness each positive or negative represents. Only the most salient, from the authors' perspective, follow below.

Among the most significant drawbacks to CUL's current separate record policy is the proliferation of catalog records for e-journals. Multiple representations of the same title create OPAC displays that can be difficult for library users and staff to interpret. CUL has tried to alleviate this situation by adding a uniform title, qualified by "Online" and the name of the aggregation, to all the machine-generated e-journal records loaded into the catalog. The library also has considered changing OPAC displays for journal title results to make better use of the qualifiers in the uniform titles. However, even with these adaptations, the displays are not ideal for catalog users. Moreover, the practice of creating a separate record for each version of an e-journal title runs counter to CONSER's current aggregator-neutral policy of creating a single record for multiple electronic versions of a title. This inhibits CUL's ability to harvest records from that

---

```
<eJrecord>
  <voyagerID>4941910</voyagerID>
  <title>Cardiovascular ultrasound (Online : Directory of Open Access Journals)
[2003-] </title>
  <relation>DOAJ</relation>
  <identifier type="URI">http://www.cardiovascularultrasound.com/home/</identifier>
</eJrecord>
```

---

**Figure 3.** An XML-encoded "Find E-journals" record

---

database and to contribute to it. Thus, while the library's policy responds to local processing demands and the needs of users for timely, accurate title and holdings coverage, it provides less than ideal displays and follows a practice that runs counter to current national serials cataloging practices.

Neither is the CUL approach entirely consistent. Despite the general policy to create separate records for print and electronic versions of journals, some instances of mulver records remain in the CUL catalog. Because of a decision the library made to purchase and load records for United States government documents from MARCIVE, and because many MARCIVE records for e-journals are mulver records, CUL elected to accept them without modification. The decision was made in the interest of expediency. While the inconsistency of this practice and the added confusion it may cause end users are unfortunate, trends in the federal government publications universe suggest that the problem may be only temporary. As the publication of government information continues to shift from tangible to electronic format, separate e-version records will likely replace the multiple version records.

Another shortcoming is that the brief, machine-generated e-journal records lack subject analysis. The absence of subject headings and classification limits subject access to title keyword searching (assuming the journal title includes subject-related words, which is not always the case). Since analysis of CUL's catalog transaction logs indicates that fewer than 6 percent of all catalog searches are subject searches, library managers do not believe that the lack of subject headings in these records will greatly compromise catalog searching, but the likelihood of serendipitous discovery, even via keyword searching, is reduced. Because the brief records carry no controlled vocabulary terms for subject access, programmatic

breakdowns of the title list by subject is rendered nearly impossible.

A further drawback to the abbreviated, machine-generated records is that they do not include any linking information to inform users about preceding or succeeding titles or other related titles, including print versions. E-journal providers vary in their treatment of title changes, but in many cases a journal can only be retrieved under its latest title, even though earlier issues may be available online. A user searching the catalog under an earlier title may not get a result. Reference librarians have been made aware of this situation and need to keep it in mind as they assist users in searching for e-journals. However, users searching without assistance from library staff well may remain ignorant of all holdings available to them.

The lack of persistent identifiers for individual e-journal titles is another compromise that the brief records necessitated. The absence of a unique, stable record identification tag makes the method resistant to services that depend on identifiers, such as bibliographic record numbers for match points in record updates. Such services include electronic resource management applications, for example, the Innovative Interfaces Electronic Resource Management system. And, while CUL staff are able to manipulate and manage large volumes of e-journal metadata quickly and efficiently, e-journals that are not issued as part of an aggregation continue to be excluded from the automated maintenance routine. Thus, no systematic refreshing of these titles takes place; cataloging staff continue to maintain these titles manually and only on an ad hoc basis.

Depending on external metadata suppliers such as Serials Solutions introduces other complexities into processing workflows: ongoing inclusion of journal titles and aggregations in vendor databases needs to be

monitored by library staff; workflows need to handle titles with diacritics in a normalized way; workflows need to account for titles with initial articles to ensure correct indexing in title browse displays; and holdings and coverage data, which Serials Solutions receives directly from publishers, is not always reliably accurate. Experience also has shown that vendors do not always provide timely information about e-journals. While expedient, reliance on external providers for e-journal-related data may result in some inaccuracies in catalog records.

Finally, CUL's automated workflows are not yet sufficiently mainstreamed for handling by lower-level library staff. The need for information technology-savvy librarians and staff members to process the routines introduces the possibility of processing bottlenecks. Maintenance of titles in both the catalog and standalone lists or databases also requires double maintenance and makes keeping both the catalog and the standalone service in sync more challenging.

Yet despite these potential shortcomings (and other possible disadvantages not specifically enumerated here), the CUL approach offers several very significant advantages. Chief among these is timeliness. Library users rely upon the accuracy and timeliness of the information they find in the catalog. This is especially true in the case of electronic resources, where verification of catalog information by examining a physical piece is not possible, and the ability to connect to a resource depends upon the accuracy of the URL in the catalog record. Invalid URLs and outdated holdings information frustrate library users and staff. Using current data from Serials Solutions to generate and maintain records means that URLs and holdings information are updated regularly and require neither a separate routing for URL checking nor any manual labor on the part of CUL staff.

Though currently handled by high-level staff, the record creation and loading process will inevitably become routine. At that time, lower-level staff will be able to perform these tasks instead of librarians. The cost of automatic record creation and maintenance is already much lower than either the cost of purchasing and maintaining complete records from an outside source, or the cost of creating and maintaining full- or core-level records in-house; using lower-level staff to run the routines will reduce costs even further.

Another plus is having the means to identify e-journals by aggregator in the catalog. This allows the e-journal data gathering processes to be greatly simplified. For example, CUL can produce a complete title list of e-journals and holdings on demand, or provide some subset of the list based on other criteria, such as supplier, publisher, or coverage dates. Library staff can analyze coverage overlaps and evaluate new aggregator packages more effectively. When coupled with improved usage statistics-gathering methods for e-journals (such as Project COUNTER data), CUL's collection management decisions can be made with better and more complete information.<sup>33</sup>

The scalability of the model and its potential for use with other kinds of resources are other strong positives. Library staff also may apply the coding combinations that identify format and aggregation to bibliographic records for locally created digital collections, monographic or serial. Records for such sets therefore can be easily extracted for batch manipulation, extraction, and sharing. The coding and extract process also allows the library the flexibility to reuse MARC data for other, non-MARC-based applications. As noted above, CUL staff already are extracting the e-journal MARC data for the Find E-journals service. Once properly coded, MARC records for other library resources potentially can

be extracted, mapped to appropriate metadata schema, and used in digital repositories for resource discovery. Thus, the library can offer users multiple avenues for accessing electronic content without investing significant staff hours in creating and maintaining multiple metadata records for those resources. At CUL, MARC records have already been extracted and mapped to various metadata schema for locally produced digital library projects.

### Recommendations for Further Study

The CUL approach assumes that title-level access and holdings data are more important to end users than other trademarks of traditional serials bibliographic control, such as subject access via a controlled vocabulary, detailed descriptive notes, and classification. The library's assumption was based in part on an informal analysis of the CUL online catalog transaction logs, which indicated very low use for both subject and call number searching. However, the data in the logs can be ambiguous, represent only a snapshot at one point in time, and may not, in any case, hold true in other institutions. A more thorough examination of user needs and expectations with regard to bibliographic records for e-journals would benefit the broader community and may reveal interesting things about the way users view the metadata that libraries present to them.

The CUL approach to maintenance also assumes that the automated processes developed could have applications to other kinds of resources and are scalable. Although internal evidence suggests that both assumptions are valid, CUL has not attempted any systematic exploration of that validity, nor have librarians and staff tried to ascertain the limits of any scalability. Further investigation into that issue might be interesting and fruitful.

Another area of potential interest

would be a study comparing the total cost of the CUL homegrown automated e-journal management solution with a simpler, but superficially more expensive method, such as purchasing MARC record data from a third party. CUL's decisions were driven in part by an inability to secure funding for the purchase of such records. However, a *post hoc* examination of the process might lead to revealing data about the actual costs incurred and how they compare to the direct expenditure required for obtaining records from a third party.

### Conclusion

Although the authors believe the CUL e-journal management process to be innovative, efficient, and effective, they also readily acknowledge the contextual nature of its appeal. In the CUL environment, a heavily automated approach is a solution that is both sustainable and scalable. Alternative paths, such as the purchase of externally supplied MARC data, were closed to the library for lack of financial or human resources. Other institutions with greater or lesser means in particular areas almost certainly would arrive at different conclusions, based on the needs and expectations of their users, public services staff, bibliographers, systems staff, and others. Thus, the solutions presented here are not necessarily intended to serve as a benchmark for all other e-journal metadata management strategies; instead, they are offered as instructive examples of what can be achieved.

### References and Notes

1. Eric Lease Morgan, "Description and Evaluation of the 'Mr. Serials' Process: Automatically Collecting, Organizing, Archiving, Indexing, and Disseminating Electronic Serials," *Serials Review* 21 (Winter 1995): 1-12.
2. Erin Stalberg, "Bibliographic Access

- to Titles in Aggregator Databases: One Library's Experience at Saint Joseph's University," *The Serials Librarian* 39, no. 4 (2001): 19–24; Wayne Morris and Lynda Thomas, "Single or Separate OPAC records for E-journals: The Glamorgan Perspective," *The Serials Librarian* 41, no. 3/4 (2002): 97–109.
3. Tina E. Chrzastowski, "E-journal Access: The Online Catalog (856 field), Web Lists, and 'The Principle of Least Effort,'" *Library Computing* 18, no. 4 (1999): 317–22; Jim Holmes, "Cataloging E-journals at the University of Texas at Austin: A Brief Overview," *The Serials Librarian* 34, no. 1/2 (1998): 171–76.
  4. Carol A. Kiehl and Edward H. Summers, "Comprehensive Access to Periodicals: A Database Solution," *Library Collections, Acquisitions & Technical Services* 24, no. 1 (2000): 33–44.
  5. Marian Shemberg and Cheryl Grossman, "Electronic Journals in Academic Libraries: A Comparison of ARL and Non-ARL Libraries," *Library Hi Tech* 17, no. 1 (1999): 26–45.
  6. Thomas R. Sanders, Helen Goldman, and Jack Fitzpatrick, "Title-Level Analytics for Journal Aggregators," *Serials Review* 26, no. 4 (2000): 18–29; Holmes, "Cataloging E-journals."
  7. Louise B. Rees and Bridget Arthur Clancy, "Cataloging Electronic Journals: Learning to Weave the Web," *Internet Reference Services Quarterly* 3, no. 3 (1998): 29–43.
  8. Karen Calloun and Bill Kara, "Aggregation or Aggravation? Optimizing Access to Full Text Journals," *ALCTS Journal Online* 11, no. 1 (2000), under "Separate Record Technique." Accessed Jan. 22, 2005, [web.archive.org/web/20010214020324/www.ala.org/alcts/alcts\\_news/v11n1/gateway\\_pap15.html](http://web.archive.org/web/20010214020324/www.ala.org/alcts/alcts_news/v11n1/gateway_pap15.html); Wayne Jones, "More Than One Record," *The Serials Librarian* 41, no. 1 (2001): 17–20.
  9. William A. Britten et al., "Access to Periodicals Holdings Information: Creating Links between Databases and the Library Catalog," *Library Collections, Acquisitions & Technical Services* 24, no. 1 (2000): 7–20; Wayne Morris and Lynda Thomas, "Single or Separate OPAC Records for E-journals."
  10. Charity K. Martin and Paul S. Hoffman, "Do We Catalog or Not? How Research Libraries Provide Bibliographic Access to Electronic Journals in Aggregated Databases," *The Serials Librarian* 43, no. 1 (2002): 61–77.
  11. *Ibid.*
  12. Library of Congress, Program for Cooperative Cataloging Standing Committee on Automation Task Group on Journals in Aggregator Databases, *Final Report*, Jan. 2000. Accessed May 20, 2004, <http://lcweb.loc.gov/catdir/pcc/aggfinal.html>.
  13. Karen Smith-Yoshimura, e-mail to RLG Technical Services Strategy Focus Group mailing list, Jan. 24, 2003.
  14. Stalberg, "Bibliographic Access to Titles in Aggregator Databases."
  15. Yiu-On Li and Shirley Leung, "Computer Cataloging of Electronic Journals in Unstable Aggregator Databases: The Hong Kong Baptist University Library Experience," *Library Resources & Technical Services* 45, no. 4 (2001): 198–211.
  16. *Ibid.*
  17. Britten et al., "Access to Periodicals Holdings Information"; Robert N. Bland, Timothy Carstens, and Mark A. Stoffan, "Automation of Aggregator Title Access with MARC Processing in Member Libraries of WNCLN," *Serials Review* 28, no. 2 (2002): 108–12.
  18. Bland, Carstens, and Stoffan, "Automation of Aggregator Title Access."
  19. Calloun and Kara, "Aggregation or Aggravation?"; Elizabeth S. Meagher and Christopher C. Brown, "Gold Rush: Integrated Access to Aggregated Journal Text through the OPAC," *Library Resources & Technical Services* 48, no. 1 (2004): 69–76.
  20. Sanders, Goldman, and Fitzpatrick, "Title-Level Analytics"; Li and Leung, "Computer Cataloging."
  21. Daniel Chudnov, Cynthia Crooker, and Kimberly Parker, "Jake: An Overview and Status Report," *Serials Review* 26, no. 4 (2000): 12–17.
  22. Meagher and Brown, "Gold Rush."
  23. Bland, Carstens, and Stoffan, "Automation of Aggregator Title Access."
  24. Stalberg, "Bibliographic Access."
  25. Li and Leung, "Computer Cataloging."
  26. Bland, Carstens, and Stoffan, "Automation of Aggregator Title Access."
  27. Terry Reese, "Aggregate Record Management in Three Clicks," *D-Lib Magazine* 9, no. 9 (2003), under "Record Overlay." Accessed Jan. 22, 2005, [www.dlib.org/dlib/september03/reese/09reese.html](http://www.dlib.org/dlib/september03/reese/09reese.html).
  28. Library of Congress CONSER Program, "FAQ on the Aggregator Neutral Record," May 29, 2003. Accessed June 15, 2004, [www.loc.gov/acq/conser/agg-neut-faq.html](http://www.loc.gov/acq/conser/agg-neut-faq.html).
  29. Library of Congress CONSER Program, "Single Records for Online Versions of Print Serials." Accessed June 15, 2004, <http://lcweb.loc.gov/acq/conser/singleonline.html>.
  30. Library of Congress, CONSER Program, "FAQ on the Aggregator Neutral Record."
  31. Terry Reese, "MARCEdit Home Page." Accessed July 23, 2004, <http://oregonstate.edu/~reese/marcedit/html>.
  32. The authors will make the SQL queries available to interested parties.
  33. Project COUNTER (Counting Online Usage of Networked Electronic Resources) is an international initiative designed to serve librarians, publishers and intermediaries by facilitating the recording and exchange of online usage statistics. More information can be found on the COUNTER Web site. Accessed Jan. 22, 2004, [www.projectcounter.org/about.html](http://www.projectcounter.org/about.html).