Commercial Services for Providing Authority Control: Outsourcing the Process

Sherry L. Vellucci
St John's University

Introduction

Throughout this conference we have discussed the importance of authority control, examined new perspectives, and looked at cutting edge projects for name, title and subject authorities within a variety of information communities. Much of the discussion focused on automated processes and data sharing. This paper will examine another side of automation, the commercial services that are available for providing authority control.

Before I begin I will say that outsourcing, whether the entire cataloging operation or only specific parts of it, is a controversial issue. A review of the existing literature shows that for every librarian satisfied with the results of outsourcing the cataloging process, there are just as many who believe that outsourcing cataloging will have negative consequences for both the library and the profession. No matter where you stand on this politically-charged spectrum, understanding the authority control services that are currently available from commercial vendors will allow you to make informed decisions.

I'd like to begin with a definition of outsourcing. The American Library Association's Outsourcing Task Force defined outsourcing as "contracting to external companies or organizations, the functions of cataloging that would otherwise be performed by library employees." The ALA Task Force viewed outsourcing as a useful management tool when the decision to outsource is based on accurate information about the local library's needs (A.L.A. Outsourcing Task Force, p. 23). The Report cites the results of several different surveys that were conducted to discover the extent to which public and academic libraries outsource the cataloging process. One survey conducted in 1998 by the Urban Libraries Council discovered that of the public libraries in metropolitan areas that were surveyed, 61% responded that they outsourced some part of the cataloging process, although none used commercial service providers for more that 50% of the overall process (A.L.A. Outsourcing Task Force, p. 24). This survey, however, did not isolate authority control from the rest of the cataloging process. Another survey, which focused only on academic libraries was conducted by Bénaud and Bordeianu a year earlier. They found that 71% of Association of Research Libraries (ARL) libraries and 56% of the non-ARL medium-sized academic libraries outsourced some portion of the cataloging process. They also discovered that larger libraries outsourced more cataloging than smaller libraries. When authority control was isolated from the overall cataloging process, Bénaud and Bordeianu report that "over half of the libraries surveyed (55%) outsource or have outsourced authority control and of the remaining 45%, an additional 14% plan to do so in the future (1998, chapter 4). All of this indicates that outsourcing cataloging in general, and the authority control process in particular, has become a common practice in both academic and public libraries in America.

So, why do librarians turn to commercial vendors to perform the authority work for their local catalog? Again, by examining the literature I found a group of core reasons that most libraries held in common.

- Authority control is labor intensive
- Staff efforts better spent elsewhere
- Controlled costs
- Increased efficiency
- Available expertise not otherwise resident in the cataloging unit
- Improved consistency and quality of the catalog database

Many factors influence these rationales, including budget cuts, staff cutbacks and the expanding need for professional librarians to cover other user services in the digital environment. Without the necessary staff, authority control is often neglected. If the required integrity of the catalog is to be maintained, many libraries must look beyond their internal resources. The good news is that this is one area where new computer applications excel in both speed and accuracy and advances made in recent technology make it significantly easier to outsource authority work (Bénaud and Bordeianu,1998).

As in the rest of the corporate world, there have been startups, buyouts, and mergers among library vendors. The United States currently has at least six agencies that provide authority control services for libraries. These include Library Technologies, Inc., OCLC's WLN MARC Record Service (MARS), MARCIVE, Follett Software, Inc., Autographics, and Internet Systems, Inc. Of these, the primary vendors for large libraries are LTI, MARS, and MARCIVE.

Types of Authority Control Services

There are several different types of authority control services offered by vendors, most of which are performed after the cataloging process is completed, because the vendor works from the bibliographic record. Vendor-supplied authority work is primarily accomplished using a batch processing method. This requires sending the databases to the vendor and often has a turn-around time of several weeks or months, depending on the size of the database. The vastly superior networking technologies now make it possible to send bibliographic records to vendors using a File Transfer Protocol (FTP). The variety of services offered includes retrospective cleanup; ongoing authority maintenance; and periodic updates.

A. Retrospective cleanup
Many libraries turn to vendors to supply their authority control after, or as part of a retrospective conversion project. Another prime time to outsource authority control is when the library migrates to a different online catalog or integrated library system. This is an ideal time to upgrade bibliographic records, update and correct access points, and provide a richer syndetic structure for the catalog.

When outsourced authority control is part of a retrospective conversion project, the authority control portion of the project is completed after the bibliographic records have been converted to machine-readable form and cleaned up. Because of the limitations of early technology and the expense of computer storage space, many libraries that converted their

bibliographic records to MARC format in the early years of migration to online catalogs did not convert and integrate their authority files at that time. In these cases, the process of converting the bibliographic records is already completed and the retrospective authority control cleanup is performed to provide current and accurate headings linked to the authority records with their accompanying references. After the authorization of the bibliographic base file is completed, a copy of the individual library's authority file is retained by the vendor as a master file. This master file can then be used in ongoing authority maintenance.

B. Ongoing Authority Control

Improved technology has made uploading and downloading records between libraries and vendors an easy process, thus making it possible to provide vendor-processed authority control on an ongoing basis. As everyone knows, authority work never ends—new headings are added, existing headings are updated, changed or deleted and new conflicts between headings can arise. The ongoing aspects of authority control fall into two categories:

- Authority control for newly cataloged bibliographic records
- Authority control of previously cataloged records where headings have changed

Using a File Transfer Protocol (FTP) or a web-based user login, vendor-provided authority control for newly cataloged bibliographic records can be executed on a monthly, weekly or daily basis, depending on the volume of bibliographic records added by the library. Several vendors now offer overnight turnaround time and one vendor (LTI) offers authorities processing using FTP in as little as one hour for files up to 10,000 records. In addition, LTI offers a 'Real-Time Authority Control" (RTAC) service based on client-server technology. RTAC allows the cataloger to send a bibliographic record during the cataloging process, have it authorized and returned to the cataloger along with all appropriate authority records almost instantly. However, the RTAC application must be integrated into local system record editors, and vendor systems have been slow to implement this integration.

In addition to ongoing authority control for newly cataloged materials, catalogers need to monitor changes to previously authorized headings. Authority vendors can provide a notification service to help with this process. The vendor tracks the updates and changes made to the national authority files used for the library's catalog and will automatically inform the library when changes are made to a heading that matches an authority record in the library's master file. New authority records for unmatched headings that were previously processed are also obtained. Most vendors allow you the option of deciding which updated authority records are included in your notification (e.g., only changes in 1xx headings, or changes in 4xx and 5xx fields, etc.) This notification can be done on a monthly, quarterly or yearly basis. The vendor will either send a replacement authority record accompanied by updated bibliographic records, or a report that lists which records have been changed, replaced or deleted so that changes can be made in-house by local staff. This will depend on the services that were contracted by the library.

D. Periodic File Reauthorizations

If the library has completed a one-time retrospective authority control database cleanup, but chooses not to contract immediately for ongoing authority control service and does not have the staff to maintain headings after the initial cleanup, periodic authorization updates for headings may be the way to go (e.g., every few years). This requires the library to export and

reauthorize the entire bibliographic database (or a selected portion). The reauthorization process is much like a retrospective cleanup, except that the library does not have to resubmit the existing authority records; only the bibliographic records are required for this process. After the database has been reauthorized, the bibliographic records are reloaded into the local system database, then the new authority records are loaded and new indexes are built.

It should be noted that all of these processes involve not only the authority control vendor, but the integrated library system (ILS) vendor as well. Maintaining good relations with both vendors is important, especially if the ILS vendor must reprogram some software as in the case of LTI's RTAC, which must be integrated into the system record editor. The library should also verify with the ILS vendor that the necessary load tables are in place to allow the processed bibliographic and authority records to be loaded back into the library's local system.

The Vended Authority Control Process

Now that we have looked at the available services that an authorities vendor can offer, let's look more closely at the process itself. The first step is to develop a project profile that lists all of the library's specifications for each step of the processing. The project manager should work closely with the vendor representative to ensure that the specific available options and the pricing structure for each part of the project are understood. Most authority vendors are flexible in customizing the process to the specific needs of the library.

The linkage rate (i.e., matched headings) usually runs from 80% to 95% depending on the libraries conformance with national cataloging standards and the type of material. Pricing structures vary from vendor to vendor and from service to service. Vendors may charge a 'per record processed" fee, a 'per heading processed" fee, or a flat rate based on database size, with additional charges for customization. It is important to have some idea of the vendor's linkage rate for the library's database before beginning the project, for the costs for manual review and in-house review and cleanup will be affected by the percentage of records that were matched during the batch processing.

A. Preprocessing
Batch processing involves a series of cleanup operations. Prior to running your database against the national authority files, vendors will preprocess your headings using a variety of software programs to normalize the headings and eliminate common errors and inconsistencies. Several vendors offer the option of cleaning up the entire bibliographic record at this time as well. The library should be able to customize these corrections to take into account local cataloging practices. Some of the most common corrections include:

- Duplicate record resolution
- Abbreviation expansion
- Correction of common errors (typographic)
- Conversion of obsolete MARC subfields
- Deletion of canceled MARC subfields
- Direct-to-indirect geographic subfield conversion (LCSH)
- Pseudonym processing
- Updated General Material Designators (GMD)

- Initial article and filing indicator validation or correction
- Correction of spacing, capitalization, punctuation

A library can elect additional services for this preprocess stage. These might include conversion of non-MARC records to MARC 21 format; provision of item level holdings conversion to accommodate local circulation systems; generation of smart barcode numbers and barcode labels; and bibliographic record enrichment such as table of contents and summary enrichment.

B. Machine match against master authority files

Once the preprocessing cleanup is finished and the headings have been normalized, the bibliographic records are run against the master authority files selected by the library, comparing authorized and variant heading forms. The vendor also maintains a supplemental authority file developed from manual review of previous customer files for authorized headings that are not included in the national files. The master authority files commonly include:

- Library of Congress Name Authority File (LCNAF)
- Library of Congress Subject Headings (LCSH)
- National Library of Medicine Medical Subject Headings (MeSH)
- National Library of Canada Canadiana
- Library of Congress Children's Subject Headings (Annotated Card Program)
- Sears Subject Heading
- Genre Headings

When a match occurs, the authority heading record is linked to the bibliographic record heading. If the heading matches a *See* reference (4xx) the authorized heading (1xx) replaces the incorrect heading in the bibliographic record. When all possible headings have been matched to an authority record, they are inserted back into the bibliographic records, replacing the existing headings in each record. The final step in this part of the process is to extract all of the linked authority records and return them along with the processed bibliographic records, to be loaded back into the library's local system. The local system then builds its index tables from the controlled headings in the bibliographic file and the *See* and *See also* references in the authority records.

Libraries have several options for transmitting their files to the vendor. Most vendors can accept and produce output in the following formats:

- 9 track magnetic tape
- 4mm data cartridges
- 8mm data cartridges
- FTP transmission via Internet

Some vendors will also supply records for all levels of a hierarchical heading for which there are matches. For example, the following heading for Beethoven's Kyrie from the Missa Solemnis would generate the following three authority records:

Heading on bibliographic record:
Beethoven, Ludwig van, ‡d 1770-1827. ‡t Missa Solemnis. ‡p Kyrie.
Authority records generated:
Beethoven, Ludwig van, ‡d 1770-1827.
Beethoven, Ludwig van, ‡d 1770-1827. ‡t Missa solemnis.
Beethoven, Ludwig van, ‡d 1770-1827. ‡t Missa solemnis. ‡p Kyrie.

C. Manual review and correction

When the machine match is completed the library must decide how to handle any remaining unlinked headings, i.e., those headings that did not match an authority record. The vendor will generate a report listing the unmatched headings, which can be used in-house for local manual review, or the library can elect to have a manual review performed by the vendor. A vendor manual review will lengthen the turnaround time for the project and can be very expensive. To determine the cost-benefit of paying for this labor-intensive review, the library should look at the percentage of headings that were linked. If the linkage rate is high and would be increased by only a few percentage points, it is probably not worth the expense (e.g., 94% linked by machine-processing plus 2% linked by manual processing). If the initial linkage rate is low, say 78%, and if an additional 15% can be linked through manual review, then manual review is probably worth the money. Therefore, the library should select a vendor that can give the highest linkage rate (90-95%) through machine processing. No vendor can provide a 100% linkage, so there will always be follow-up review and corrections on any processing project. The library should be prepared to allocate cataloging staff time to this type of clean-up when the records are returned.

Several different types of problems might be discovered by manual review. These include:

- Obvious typographical errors in names and subject headings
- Invalid terms in subject heading subfields
- Incorrect order of multiple surname and corporate heading elements
- Records that match multiple authority records
- Split headings

Mislinked headings will not be picked up by manual review, since linked headings, whether correct or not, are not reviewed manually. When the manual review is completed and corrections made, the bibliographic records with those headings are processed again against the national authority files to generate additional matches.

D. Reports

The reports provided by the vendor can be customized to the library's needs. They usually include the following lists of information:

- o Full match on 1xx field
- o Full match on 4xx field
- o Unmatched headings
- o Partially matched headings
- o Headings that match multiple authority records
- o Headings that have split
- o Incorrectly used headings

Statistical reports can provide a database profile, which analyzes the data and reports on the characteristics of the file. Pre- and post-processing data will give the library information that is useful for predicting future costs and mass storage requirements. Other types of reports can include:

- Frequency counts for each material format type (music, AV, etc.)
- Holding library information
- Average size of a record
- Number of characters in largest & smallest records

- Number of fields per record summary
- Field use summary
    1. Number of times a field occurs
    2. Average number of times a field is used per record
    3. Number and percentage of records in which field appears
    4. Average length of each tag

In addition to these common reports the library should be able to request customized reports to meet its specific needs.

Future prospects

This presentation on commercial services available for authority control provides background information for outsourcing and gives an overview of the process. I have focused on the vendors and processes that are now available in the United States and Canada, but I do not claim to know the state of commercially provided authority control here in Europe and elsewhere in the world. I do know that there are serious issues that must be worked out before authority control vendors can cope with the types of international authority control that we have discussed during this conference. Before we can begin to approach commercial processing using a virtual international authority file, the FRANAR issue needs to be resolved. Many interoperability problems must be solved, including issues of MARC/UNIMARC authority format harmonization, and methods that will allow libraries to use a variety of national authority databases for the matching process, and provide the flexibility to select the authorized form that is appropriate for the local users. Fortunately, interoperability issues are at the forefront of both researchers' and cataloging agencies' concerns and this can only help address many of the authority control issues. It may be that some of the authority control interoperability problems will be handled at the local level, but the commercial vendors must still play a major role. Perhaps the largest hurdle to overcome may be convincing vendors that authority control in an international arena is quite possibly the way of the future. One vendor with whom I spoke said that the concept of an international authority record had a certain "academic appeal," but based on historical and current world events he did not see evidence that a one-world vision was particularly successful. The idea of international authority control was of no concern to him and it was obvious that he was unaware of the work that has been going on among the national libraries and within IFLA. Fortunately not all vendors feel this way, for they are a critical partner in promoting and implementing automated authority control on the international level.

References

American Library Association. Outsourcing Task Force. (2000). *Impact of Outsourcing and Privatization on Library Services and Management.* Robert S. Martin, Principal Investigator. Texas Woman's University School of Library and Information Studies. Retrieved January 4, 2003 from http://www.ala.org/alaorg/ors/outsourcing/

Bénaud, Claire-Lise and Sever Bordeianu. (1998). *Outsourcing Library Operations in Academic Libraries: An Overview of Issues and Outcomes.* Englewood Colo.: Libraries Unlimited.

Lam, Vinh-The. (2001). "Outsourcing Authority Control: Experience of the University of Saskatchewan Libraries." *Cataloging & Classification Quarterly* 34 (2): 53-69.

Tsui, Susan L. and Carole F. Hinders. (1998). "Cost-Effectiveness and Benefits of Outsourcing Authority Control." *Cataloging & Classification Quarterly* 26 (4): 43-61.