Using CRDP to Manage Cataloging in a Federal Depository Library

by

Edith K. Beckett, Ph.D.,
https://orcid.org/0000-0003-2349-4157
Head of Technical Services, New Jersey State Library

Abstract
The New Jersey State Library has been a federal depository library since 1895. Prior to 2009, staffing limitations in the library meant that only 60 to 70 percent of the materials received from the Government Publishing Office were added to the library’s catalog. The library also had no effective practices for cataloging the growing number of digital documents in its federal depository library collection. In March of 2010, the New Jersey State Library began participating in the Government Publishing Office’s Cataloging Record Distribution Program (CRDP). Through this program, MARCIVE, a U.S.-based library services company, supplies the library with MARC formatted bibliographic records for all the new titles published by agencies in the library’s depository profile. As a result, all the materials received through the depository library program since 2010, including those in digital format, are routinely added to the library’s catalog.

Introduction
The New Jersey State Library (NJSLS), in operation since 1796, is one of the oldest state library agencies in the United States. It is in Trenton, New Jersey, and provides a full range of library-related services to state government and the citizens of the state. The technical services department at NJSLS has a staff of seven full time employees, three of which are full-time catalogers. This department is responsible for all copy and original cataloging for the general book collection and nine special collections. The technical services department is also responsible for library acquisitions, management of NJSLS’s Horizon integrated library system, including the database authority files, and bindery and preservation activities. NJSLS is also a depository library for state and federal government publications. This chapter will describe the New Jersey State Library’s efforts to provide catalog access to its tangible and electronic federal government documents using records provided by MARCIVE, Inc. as a part of the Cataloging Records Distribution Program (CRDP).

Organization Background
The services offered by the New Jersey State Library fall into three broad categories: services for state government, services for public libraries, and services for the public. Services for state government include access to on-site research collections, including extensive law, state and federal documents, and genealogy collections. The library also provides comprehensive reference and research support including access to approximately 227 databases, interlibrary loan
services, and a collection of more than one million books, documents, and journals. State government workers are also provided services through branch libraries located in the Department of Banking and Insurance, the Department of Environmental Protection, and the Department of Transportation. The library at the state Attorney General’s Office, while not a branch library, shares NJSL’s integrated library system and catalog.

Services for public libraries include the provision of broadband connectivity and related services, a statewide interlibrary loan system, and training for public library trustees and new public library directors. The State Library also administers New Jersey library laws and the per capita state aid program for libraries. The public is provided on-site access to all the collections, and New Jersey residents are permitted borrowing privileges from the circulating collections. Services are also provided to people with print disabilities through the Talking Book and Braille Center. In July 1996, the New Jersey State Library became an affiliate of Thomas Edison State College, now Thomas Edison State University. Under this arrangement the university provides administrative oversight to the State Library, but the State Library’s operations are primarily funded by appropriations through the state budget. In FY 2017 the direct funding for the State Library was $5.3 million.1

Setting the Stage

The Government Publishing Office (GPO) was established in 1861 as the primary printer for U.S. Government publications. “All printing for the Congress, the Executive Office, and the Judiciary–except for the Supreme Court of the United States–and for every executive department, independent office, and establishment of the government is required to be done at or contracted by GPO.”2 GPO “is also responsible for the acquisition, classification, dissemination, and bibliographic control of tangible and electronic government information products.”3 In December 2014, the Government Printing Office officially became the Government Publishing Office in recognition of the fact that dissemination of digital content continues to be a significant part of the organization’s efforts.4 These information products are distributed by the Superintendent of Documents through a nationwide system of over 1,150 depository libraries and an online bookstore.5

As one of the world’s largest content creators, the United States government uses the GPO and the Federal Depository Library Program (FDLP) to distribute “certain classes of Government documents free of cost to designated libraries throughout the United States and its territories. These libraries are known as federal depository libraries. Federal depository libraries must offer free, public access to their Federal collections.”6 Federal depository libraries can be regional depository libraries, that receive every publication distributed by the Government Publishing Office (GPO), or a library can be a selective depository, which means they receive publications in self-selected categories. The New Jersey State Library has been a selective federal depository library since 1895. As a selective depository, the library receives materials that support its primary mission of providing support for the work of New Jersey’s executive, legislative, and judicial branches of government. Most of the materials received through the
federal depository library program are a part of the U.S. documents collection, however appropriate titles are also added to the law collection. The federal depository library materials come in a variety of formats, including print, microform, and video. There are also computer files on various media ranging from 3.5-inch floppy disks to DVDs. Most of the materials received by the New Jersey State Library are in the areas of law, health, education, and the environment. There is also an extensive collection of Census materials, some of which dates from the 1790s, and Congressional reports and documents dating from 1776-1889 and 1921 to the present.

Providing bibliographic access to government documents has long been a concern of libraries. Government documents can be “bibliographically complex entities” due to the changeable nature of the government entities responsible for their production. In their 1987 study, researchers Turner and Latta found that a majority of the research libraries reported that many of the federal publications in their collections were “uncataloged and housed in a centralized document collection . . . [and] anticipated use, subject, and anticipated research or permanent value [were] prominent criteria used to determine the level of bibliographic control for an individual government publication.”

Lynch and Lasater suggest that while government publications are valuable sources of information, working with them in large numbers often means that they seldom receive the full cataloging afforded to other library materials. The experiences at NJSL conform to these reported findings. While the goal of the technical services department has always been to have all the materials in all the library’s collections represented in the catalog, the sheer volume of materials received through the federal depository library program has meant that staff was only moderately successful for the depository library materials. NJSL’s experiences have been much like those reported by Western Kentucky University, where this “goal” would be more accurately described as “a heartfelt commitment” rather than an achievable goal.

Cataloging duties for materials received through the federal depository libraries program are currently allotted 1/4 to 1/3 FTE of staff time. In the past, NJSL catalogers focused primarily on cataloging documents that were sent to the technical services department by the librarian responsible for the management of the federal depository collection. Most of these were documents in print and VHS/CD-ROM/DVD formats. Selected microform materials were also cataloged, but most of the materials in this format remained un-cataloged. The New Jersey State Library is an OCLC member, and the catalogers would search the OCLC database for records that matched each document sent for cataloging. Once an appropriate record was found, catalogers would make edits following local practice guidelines, and export the record into the library catalog. On average, it took about 15 to 20 minutes of staff time to get an item cataloged and ready to circulate. According to the librarian managing the federal depository documents collection, 30 to 40 percent of this collection was still un-cataloged as of 2009. The number of un-cataloged federal documents in the collection continued to increase as GPO began distributing more and more publications in digital formats. Absent specific policies or guidelines, these digital documents were treated in the same manner as microfiche documents, i.e. some of
them were selected for cataloging, but many of them were not. The current budget environment meant that adding additional staff to address the growing influx of federal documents was not a viable option; this problem had to be addressed using existing staff. Without significant changes in cataloging practices, it seemed very likely that the un-cataloged portions of NJSL’s federal depository library collection would continue to rapidly increase.

The Cataloging Records Distribution Program (CRDP)

In 2009, the librarian managing the federal depository documents collection suggested that NJSL could benefit from participation in a new program being piloted by Federal Depository Library Program (FDLP). This program, originally called the cataloging record distribution project, would provide participating depository libraries with MARC records for federal materials free of charge. The provided records would be created by GPO, and would be distributed to participants by MARCIVE, Inc. MARCIVE, Inc. is a U.S.-based library services company, established in 1981, that specializes in supplying MARC bibliographic and authority record services to libraries.12 The entry for CRDP in the Federal Depository Library Program News & Events blog provides detailed information about the program.13 This blog posting lists the following advantages to participants:

- The service is available at no cost to participants
- Monthly delivery of quality bibliographic records
- Bibliographic records match participants’ FDLP item selection profiles, and the delivered bibliographic records can easily be adjusted to accommodate changes in the profile
- Records can be easily acquired by the participants
- Participation in the program will reduce staff cataloging time
- Participants can request community support from other CRDP participating libraries

The program is open to all Federal depository libraries on a “first come, first, served” basis. Libraries usually join at the beginning of the Federal fiscal year, but if a “program vacancy arises during the fiscal year, the next library on the waiting list is able to join the program at that time.” The program has expanded every year, and as of May 2017, was providing bibliographic records to 165 Federal depository libraries.14

To remain an active participant, libraries are required to:

- Maintain CRDP point(s) of contact with both GPO and MARCIVE, Inc.
- Retrieve, review, and load the records on a timely basis
- Profile to receive at least ten records each month, on average
- Complete an annual GPO survey about the CRDP
- Have an integrated library system which supports Resource Description and Access (RDA) bibliographic records.15

Participating libraries are also expected to make all records available in their ILS/OPAC. Libraries may edit bibliographic records according to local protocols, and there are limited provisions that allow libraries to request that MARCIVE re-post records if staff at the library fail to retrieve records within the designated one-month window. Each participant may choose to
receive all newly created bibliographic records, or all newly created and changed bibliographic records. In addition, libraries can add Historic Shelflist records to their selections, or they can elect to receive only Historic Shelflist records.

Implementation

In 2010, NJSL began participating in CRDP. Staff at NJSL elected to receive both newly created and changed bibliographic records. The new records would be edited, if needed, to make them compliant with cataloging standards, which is currently RDA, and all would contain Library of Congress subject headings. Catalogers would add the newly created records to the library’s catalog, and update OCLC to indicate that NJSL held the titles.

Many of the modifications to the records in the changed record file did not appear to affect the access points for the bibliographic items. There was also concern that batch loading the changed record files would overlay existing records and wipe out any local record modifications. This overlay problem could be eliminated by having cataloging staff review the file, but there was insufficient cataloger time to review both the new and changed record files. After some discussion, it was decided that the federal depository collection librarian would review the changed records to see if any of the changes were significant enough to require replacement of the existing NJSL catalog records. Those records with significant changes would be flagged for the cataloger, who would edit the records to meet local specifications and upload them into the library’s catalog. NJSL elected not to receive the records for the Historic Shelflist. The rationale for this decision was that there was not enough staff time to process new records, review changed records and deal with the records from the Historic Shelflist. NJSL’s participation in the CRDP pilot project began in March 2010 with the download of two batches of test records. The first batch of actual records, downloaded from the MARCIVE website on March 30, 2010, consisted of records for 172 tangible (print/microform/CD-ROM/DVD) titles and 466 digital titles.

The new records for NJSL are uploaded to the MARCIVE website each month as a series of four files, each containing varying numbers of records. The number of records in each file is determined by the documents cataloged or modified by GPO each month that fit NJSL’s depository selection profile. There are new and changed record files for items issued in tangible formats, and the same for items issued in digital format. All the digital records have persistent uniform resource locators or PURLs. “The FDLP uses persistent uniform resource locators (PURLs) to provide stable URLs to online federal information. When a user clicks on a PURL, the request is routed to the federal publication. As federal agencies redesign and remove information from their sites, GPO staff reroute PURL entries to the appropriate location.”

The records for items in tangible formats do not have PURLs, but many have fields indicating that the title was issued in additional formats. Staff at NJSL decided that adding PURLs to the records for tangible items would be beneficial to users. It was also decided that to facilitate resource sharing, our OCLC holdings would be updated for all the records added to our local catalog. Once NJSL staff decided how they wanted to use the records, understood how the
records were structured on the MARCIVE website, and worked out the procedures for retrieving and storing the records, the next steps in the process were figuring out the most effective way to edit and process the records.

Creating an Effective Workflow

One of the earliest decisions faced by the staff was how to edit the records supplied by MARCIVE to conform to NJSL local protocols. Staff decided to use the MarcEdit software, and batch editing to process the records. MarcEdit is a MARC editing utility designed to support all MARC formats, including MARCXML. It is available as open source software and was developed by Terry Reese at Oregon State University.19 Several members of the technical services staff attended a MarcEdit training workshop in 2009, and so were familiar with the basic functions of the software. Additional training was provided by online tutorials.

The cataloger uses the program’s MarcBreaker to convert the records downloaded from MARCIVE into a human-readable mnemonic format. Then, the MarcEditor is used to review the records, and perform batch and individual edits. Next, the MarcValidator is used to check the structure of the records and alert the catalogers to issues with field/subfield structure that could cause problems when the records are imported into the library catalog. Then, the edited and validated records are converted back into the MARC format using MarcMaker.20 Finally, OCLC holdings are updated for each record using batch processing, and batch import functions are used to add the edited records to the library’s catalog.

CRDP was the New Jersey State Library’s first experience using batch uploads of vendor-supplied records. Prior to CRDP, every record added to the catalog was created or selected by a cataloger, and selected records were individually reviewed or edited prior to being uploaded into the catalog. NJSL management wanted to retain this level of quality control, so the decision was made to review and if needed edit every CRDP record prior to upload. This decision created a couple of initial challenges. First, the sheer volume of records, an average of 600 records each month, meant that the cataloger for the federal depository collection would have to devote significant time to this process. Since only 1/3 to 1/4 FTE of cataloger time was allotted for this collection, an extremely efficient process would have to be put in place to meet these editing needs. Working at this level of intensity also made cataloger fatigue a real possibility. Since a single person needed to review and edit all the CRDP records, procedures had to be developed to guard against cataloger burnout, assess the quality of the uploaded records, and reduce errors caused by cataloger fatigue. The solution implemented by NJSL staff involved several parts: working with small batches of records, working in short editing sessions, and keeping a written log to track the process.

After some trial and error, it was determined that fifteen records were the optimal batch size. This means that larger files are easily divided into workable units using MarcEdit’s MarcSplit function, e.g. the 662 e-docs records from the November 2016 download was processed as 44 fifteen-record files and a single two record file. These smaller files are then edited over the course of a month. Files of records for tangible items are generally smaller, e.g.
the November 2016 download contained 25 records. The cataloger typically spends eight hours each week editing records. This usually results in an average of 360 records edited each week. The editing sessions are usually no longer than two hours, and an average of six files, of 15 records each, are edited during each session. Staff has found that this time and pace is fast enough to keep up with the flow of records without creating undue fatigue.

Staff at NJSL decided that, when available, PURLs would be added to the records for tangible items. This decision was taken to support user access by ensuring that whenever possible, users have direct access to documents through the catalog. This step requires the cataloger to locate the OCLC record for the electronic version of the tangible item and copy the PURL into the record being edited. This process is aided by GPO catalogers’ decision to include MARC field 776 in many of the records. This field indicates what additional formats are available, e.g. microform, online, etc., and gives some descriptive information about that additional format, including the OCLC control number of a bibliographic record for that related format. Approximately 95 percent of the tangible items selected by NJSL have PURLs that are added to the appropriate bibliographic records.

The cataloger for the federal depository collection also uses a written log to track progress through each batch of records. The cataloger records the OCLC Control Number for each record edited, whether there is an electronic version and a PURL available for tangible items, whether OCLC has been updated to indicate NJSL holdings, and whether the record has been added to the Horizon ILS. This written log provides a mechanism to maintain workflow consistency over multiple files and over many working sessions. The log easily allows the cataloger to see what batch is currently being edited (the latest entry in the log), and what processing has been completed on the files. The log also facilitates OCLC batch processing by providing a list of OCLC control numbers that have been edited. The OCLC control numbers from the log are also used for random quality assurance spot checks in the Horizon catalog. The written logs are created from standard 5-subject college-ruled spiral notebooks, with hand-ruled columns to track the work completed on each batch of records.

**Importing CRDP Records**

Once the editing process was established, the next challenges were related to getting the records into the Horizon catalog. The New Jersey State Library uses import profiles to automatically create item records for each newly added bibliographic record. For other collections, catalogers generally work with smaller batches of ten to thirty individually selected records. Once these small batches have been imported into the library’s catalog, each automatically created item is manually modified with specific barcode and collection information. This was impractical for the hundreds of records in the CRDP batch imports, so two record import profiles were created to minimize the processing work for the CRDP records. The import profile for electronic records automatically indicates that the item is in electronic format, is available in the U.S. Documents collection, and assigns “online” as the item’s location.
The import profile for tangible item records was a little more challenging. In cases where the item had already been cataloged, and the CRDP used the same OCLC record as NJSL catalogers, the imported CRDP record would simply overlay the existing catalog record, and no additional staff intervention was needed. However, quite a few of the tangible records were for titles that had never been cataloged. Since these records could be matched with physical items in the library’s collection, the initial plan was to retrieve these items and update the catalog records with the appropriate barcodes and locations. This required an overwhelming amount of staff time to retrieve, process, and re-shelve these items. The final implemented solution uses the import profile to designate that the items are available in the collection. The result is that existing items and their barcodes remain unchanged when the new CRDP bibliographic record is imported to overlay an existing bibliographic record, but if there are no previously existing items/bibliographic records, new items are created with a status of “Ask Librarian.” Reference and access services staff have received training to know that if someone looks for an item with this status, it is most likely on the shelves filed under the Superintendent of Documents classification number. If the item is needed for circulation, a barcode is added to the record. However, if the item is only used in-house, it generally remains un-barcoded. Resolving this issue removed the last implementation hurdle, and the cataloger for the federal depository collection began to successfully add hundreds of records for documents to the catalog each month.

Quality Control, Both Before and After Import

One of the biggest concerns for NJSL management was ensuring that the vendor-supplied records met the library’s quality benchmarks. These previously established benchmarks require full level descriptive catalog records, created in accordance with current cataloging standards, e.g. RDA, that have at least one name or subject access point, and contain only subject access points based on valid Library of Congress Subject Headings (LCSH). The records should not contain locally excluded subfields and should include only name access points created in accordance with Library of Congress Name Authority Cooperative Program’s (NACO) rules and recommendations. Staff at NJSL ensures the CRDP records meet these benchmarks with multiple levels of staff review. The first level of review is performed by the cataloger who examines each record and edits them as needed to meet the desired benchmarks. The second level of review occurs when other technical services staff look specifically at the new name and subject authority records generated by the record imports. Once the records have been imported into the library’s catalog, the third level of review is conducted by the depository collection cataloger, and by the librarian who manages the federal depository collection. These librarians randomly select records from each month’s imports for review.
CRDP and Local Authority Records

The primary challenge related to CRDP participation is local authority control for the new records. The New Jersey State Library maintains local authority files for its Horizon catalog, and batch imports of hundreds of records can often generate hundreds of new authority records. For example, a recent import of 120 bibliographic records generated 45 new name authority records and 141 new subject authority records. Technical services staff checks each new authority record to verify that it is not an alternative form of a name access point already represented in the catalog, or to confirm that the subject is a validly constructed LCSH access point. New authority records are generated by all catalogers, so the CRDP generated authority records are often supplemented by an additional ten to fifty records added during a typical day. Checking hundreds of authority records each day proved to be an overwhelming task. After some discussion, it was decided to limit the number of records in each CRDP import batch to no more than 120, and to limit the large imports to no more than twice each week. Breaking the editing and processing into stages, and tracking progress on the stages, allows many independent editing/upload sessions. This reduces cataloger fatigue, minimizes the chances of error, and allows gradual integration of large numbers of bibliographic records and their related authority files into the library catalog.

CRDP Benefits for Technical Services

Work with CRDP has helped raise the profile of the technical services department within the library. Working closely with the federal depository collection librarian has become a model for the way that catalogers interact with the librarians who manage the other special collections. As an example, catalogers now routinely send lists of titles to the special collection managers when new materials are cataloged for their collections. This is beneficial, because there can be a significant lag between placing an order for materials, and the receipt of those materials for cataloging. Sending the notifications lets the collection managers know that the ordered materials are now available, provides another level of review of the records, and makes the work of technical services much more visible to library colleagues outside of the department. Catalogers have also become more aware of the relationship between their work as catalogers and the functionality of the library catalog authority files. This results in fewer name and subject heading errors in the bibliographic records. Catalogers have also begun work on cataloging and preserving legacy materials in the federal depository collection. This was enabled in part because NJSL now receives primarily electronic documents in this collection, and no additional processing is required beyond the work with CRDP. In addition to the increased productivity from participating in CRDP, the technical services department has also benefited from the expertise that has developed because catalogers are consistently working with large numbers of good quality catalog records.

CRDP was the New Jersey State Library’s first experience using large batches of vendor supplied bibliographic records. One of the unanticipated benefits of participation in CRDP is a
willingness to use vendor-supplied bibliographic records for other collections. Most notably, cataloging staff used MarcEdit and vendor-supplied bibliographic records to eliminate a cataloging backlog of approximately 4,000 records for the general e-book collection. Currently, catalog records for all new titles added to the general e-book collection, as well as some titles in the Reference collection, are created using vendor-supplied records.

NJSL plans to continue participation in CRDP for the foreseeable future. Some future includes developing facility with regular expressions to further streamline some editing tasks in MarcEdit. It would also be useful to collect data to see if there is empirical evidence of increased use of federal depository materials. Staff may also explore whether adding fee-based services from MARCIVE, e.g. purchasing authority records, could further streamline implementation.

Conclusion

CRDP implementation has been extremely beneficial for the New Jersey State Library. Table 1 shows the total numbers of federal depository documents cataloged during the first seven years of NJSL’s participation in the program.

Table 1: Federal Depository documents cataloged each year of CRDP participation

<table>
<thead>
<tr>
<th>Year</th>
<th>Number of records cataloged</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010</td>
<td>7081</td>
</tr>
<tr>
<td>2011*</td>
<td>5230</td>
</tr>
<tr>
<td>2012</td>
<td>7289</td>
</tr>
<tr>
<td>2013</td>
<td>7257</td>
</tr>
<tr>
<td>2014</td>
<td>6344</td>
</tr>
<tr>
<td>2015</td>
<td>7062</td>
</tr>
<tr>
<td>2016</td>
<td>8344</td>
</tr>
</tbody>
</table>

* The lower number of records cataloged in 2011 is a reflection of staffing changes, and staff time spent away from cataloging working on special projects.

Given the available staff, it is extremely unlikely that cataloging many of these 48,607 records could have been completed without CRDP. In 2013, physical space constraints prompted the decision to increase the number of federal depository library program documents received in digital formats; receipt of catalog records through CRDP was a major factor in support of this decision. CRDP has resolved many of the cataloging issues for this collection, but NJSL will need to assess the impacts of this change, and develop strategies to preserve and provide access to this digital content. For the New Jersey State Library, participation in CRDP, along with the use of MarcEdit and OCLC’s batch processing has resulted in significantly more records for federal depository materials in the collection, getting us closer to that elusive goal of having all of the materials in all of the collections represented in the library catalog. It has also resulted in improved service by the technical services department to both library staff and the users that they service.


8 Turner and Latta, *Current Approaches*, 19.


13 Federal Depository Library Program, "Cataloging Record Distribution Program," 
FDLP News & Events (blog), entry posted October 24, 2012, accessed October 10, 2017, 

14 “Cataloging Record Distribution Program,” last modified October 24, 2012, accessed 

15 "Cataloging Record."

16 The GPO Historic Shelflist is the U.S. Government Publishing Office's (GPO) historic 
card catalog of over one million 3x5 cards arranged in Superintendent of Documents (SuDoc) 
additional information.

17 “RDA: Resource Description and Access was developed by the RDA Steering 
Committee (formerly the Joint Steering Committee in for the Development of RDA) as part of its 
strategic plan (2005-2009) to replace the Anglo-American Cataloguing Rules, 2nd Edition 
Revised, which were first published in 1978. RDA is a package of data elements, guidelines, and 
instructions for creating library and cultural heritage resource metadata that are well-formed 
according to international models for user-focused linked data applications” RDA Steering 
Committee, "About," RDA Toolkit, accessed September 15, 2017, 
http://www.rdatoolkit.org/node/44.

18 “Federal Depository Library Persistent URL Home Page," accessed August 2, 2017, 

19 MarcEdit Development Home Page accessed August 2, 2017, 
http://marcedit.reeset.net/.

20 "Features -- MarcEdit Development," accessed August 8, 2017, 
http://marcedit.reeset.net/features.

21 Atifa Rawan, Cheryl Knott Malone, and Laura J. Bender, “Assessing the Virtual 
710-726.