Analysis of Readex's Serial Set MARC Records: Improving the Data for the Library Catalog

By Daniel Draper[a] and Naomi Lederer[b]

[a]Metadata and Preservation Services, Morgan Library, Colorado State University, Fort Collins, CO 80523-1019 USA Daniel.Draper@colostate.edu 1-970- 491-5272 FAX 1-970-491-4611
[b]College Liaisons, Morgan Library, Colorado State University, Fort Collins, CO 80523-1019 USA Naomi.Lederer@colostate.edu

Abstract

Colorado State University Libraries (CSUL) purchased the digitized *United States Congressional Serial Set*, 1817-1994 and *American State Papers* (1789-1838) from the Readex Division of NewsBank, inc and, once funds and records were available, the accompanying MARC records. The breadth of information found in the *Serial Set* is described, along with the difficulties in using the print version (incorporated in the literature review, which includes citations of announcements of the digital collections and reviews of the software). The digital version of the *Serial Set* has its advantages, but there are additional rewards (much greater discovery opportunities) when items in the digital collection are directly accessible from the library catalog. The purchased MARC records, while overall excellent, had problems that needed to be corrected before they were loaded into CSUL's Innovative Interfaces library catalog. Patron access impact was used as a criterion when determining which of the records would be fixed before loading. High impact problems were identified and solutions derived for: multiple 245 (title) fields; 245 second indicator zero with titles beginning with a, an, or the; dollar sign used in text; fixed field date; languages; subject headings; creating proxy URLs; classification numbers; and author authority control (e.g. corporate entries and presidential entry errors).

1. Introduction

Colorado State University Libraries (CSUL) recently purchased the digitized *United States Congressional Serial Set*, 1817-1994 and *American State Papers* (1789-1838) from the Readex Division of NewsBank, Incorporated. This collection encompasses an extensive compilation of 15,579 physical volumes (consisting of 14,277 numbered *Serial Set* entries, including multi-volume sets) with 11,597,235 pages. The *American State Papers* contains 6,354 publications, in 38 volumes with 40,389 pages. The 370,175 individual documents in the *U.S. Congressional Serial Set* include House and Senate documents, reports, and journals. There are over 71,000

maps (A. A. Imholtz, Jr. personal communication, June 21, 2011). The House and Senate contracted with free enterprise printers to publish all documents until 1861, when the Government Printing Office began production (Harrison, 2010, 22.). The wide range of topics found in the *Serial Set* go well beyond strictly Congressional. There are government agency and special reports as well as presidential (and therefore executive branch in origin) communiqués. *Serial Set* contents go well beyond the functions of the government and delve into historical records in archaeological digs, biographies, creations of parks and monuments, petitions and claims of widows, shipping concerns, Native American claims regarding depredations of property, geological (there are many USGS surveys and other reports), agriculture, transportation, information on indigenous tribes, and much more. "The *U.S. Congressional Serial Set* is one of the undisputed treasures of government document collections. It is filled with a depth and richness of American and world history that make it the crown jewel of any collection of government information" (Rossmann, 2004). The digital end product revolutionizes access to this preeminent collection with OCR (optical character recognition) accuracy and full text searching; the latter results in enhanced retrieval that can, however, create problems if not constructed carefully.

Digital access opens this extensive primary source collection to a wider audience because it is easier to use; instead of using a printed index and then, in the case of researchers at Colorado State University, having to request pertinent volume(s) from storage (delay of delivery of 5 to 72 hours), affiliated researchers have access to the entire collection 24/7. However, even with the capability to search the collections within the separate Readex databases, it is advantageous to be able to access the collection from within the library's online catalog. First of all, any number of

researchers may not even think of searching the separate database collections. Only those already familiar with government documents are likely to think of searching them. In addition, even those researchers who would seek out Congressional materials may not realize the scope of the collections. Having these items listed in the library catalog means that keyword searchers (and sometimes subject searchers--see section 5.6 for why subjects do not always work) will happen upon valuable resources that they may not otherwise discover. For example, it probably would not occur to a researcher to look within the *Serials Set* for European labor statistics from the 1880s. Researchers may not be familiar with different spellings of people's names or the multitudes of agency realignments (and renaming of) that have taken place over the years; proper (authority) headings assist in collating related items. Furthermore, the new catalog records give visibility to older and obscure materials which are still useful to researchers. Because these types of documents may not be actively sought, an additional access point improves serendipitous discovery. Moreover, purchasing an electronic *Serial Set* is a significant financial commitment on the part of the library, and any way to make the documents visible should be considered. MARC records offer powerful indexing capabilities that strengthen the ability to identify useful materials by added authors, titles, series titles, and subjects. This article provides answers to specific concerns that need to be addressed when adding Readex-provided records to a library's catalog.

Broad subjects within the Readex *Serial Set* and the *American State Papers* databases are listed on their initial Web pages as: Armed Forces and Conflicts; Discovery and Exploration (*Serial Set* only); Economics; Education; Environment, Energy and Natural Resources; Food and Agriculture; Geographic Names; Government and Politics; Groups of People, Professions, etc.;

Health; Indian Affairs; Infrastructure, Transportation, etc.; International Affairs; Legal System; Names of Acts; Personal Names; Publishing and Communication; Safety, Accidents and Disasters; Science and Technology; and Social Issues. These categories, which have numerous sub-categories, are helpful to those using the databases directly (which also allows keyword and field searches); however, many researchers will not think to use the databases, so an additional access point is desirable.

Policy is often determined on precedent and examining earlier governmental sources can help support or rebut current thinking. Reminder of the existence of the *Serial Set* encourages more thorough and comprehensive research by policymakers and their staff members (some of whom may be interns who are replaced on a cyclical basis). Additional access points, in this case via a library catalog, make it more likely that relevant material will be discovered in a timely fashion by those who are looking for earlier sources.

At CSUL, use of the *Serial Set* visibly increased once the document records for it were added to the library's catalog. The year before the MARC records for the *U.S. Congressional Serial Set* were added (June 2007-May 2008), there were 134 searches in the database and 510 document views (these are individual documents that were viewed, not number of pages viewed within). In the years after the catalog records were added there were: 140 searches (2008-9) and 2284 document views; 138 searches (2009-10) and 10,034 document views; and 137 searches (2010-11) and 7216 document views. (Calendar year totals were: 111 and 8529 in 2009; 146 and 7514 in 2010; and 282 and 5739 in 2011.) So while the number of searches in the database stayed relatively low, the number of documents viewed increased dramatically. The *American State*

Papers was searched once in the year before the records were added (August 2009-July 2010) with two document views. The following year (2010-11) there were 77 searches and 656 document views. (Calendar year total for 2011 through mid-December was 101 and 661.) This data strongly suggests that the catalog records have made a positive difference in the use of the *Serial Set*.

Direct searches of the database remained relatively consistent with slight increases that might very well be from users who discovered the resource via the record in the library catalog. Once discovered, alert researchers are likely to go directly to the *Serial Set*. Drops and increases in views may be related to course assignments. However, overall direct searches of the databases have remained stable, while use of documents has greatly increased, which to CSUL justifies the inclusion of the records in the library catalog to enhance use of this major investment.

2. Literature Review

There are currently no articles that specifically address the modifications needed when adding Readex's catalog records for the *U.S. Congressional Serial Set* or the *American State Papers*. The catalog records have only been for sale for a few years, so multiple commentaries would not be expected. In this case, the records were added to an Innovative Interfaces catalog. All of the articles listed here have endnotes, bibliographies, and/or literature reviews of interest; researchers desiring additional information will want to investigate those resources in detail.

2.1 On the *Serial Set* Itself

While as of this writing there are no articles on the topic of loading vendor-provided *Serial Set* records, the literature does contain some worthwhile related articles, which will be outlined here. (In some cases within this literature review, "*Serial Set*" will refer to both the *American State Papers* and the *U.S. Congressional Serial Set*.)

August A. Imholtz Jr. (2008), an expert on the *Serial Set*, wrote an interesting article published in *DttP: Documents to the People* about how items were selected or omitted from the *American State Papers*. He reiterates from his earlier article (Imholtz, 2003) that when Congress first began there was no uniform standard for numbering, size, distribution, or editorial control of Congressional documents. When a numbering system was finally introduced it was not retroactive. Documents from the first fourteen Congresses did not exist as a complete set. General William Hickey was the chief compiler of documents and was unable to republish every existing document in the *American State Papers*. Imholtz's article provides examples of missing items and posits reasons why many reports and documents went missing, along with a plea to track down missing items.

An article by Suzanne DeLong titled "What is in the *United States Serial Set*?" (1996) provides a detailed overview of the *Serial Set* contents. Starting with the news that use of the *Serial Set* accounts for a minimal percentage of U.S. government documents use, DeLong wrote that if people knew what was in it, they would be more likely to use the *Serial Set* in their research. She then outlines its contents and provides details of the various series (Senate Journals, Senate Documents, House Journals, and House Documents; Reports were included later). The

numbering system began with one, for the 15th Congress. Delong then closely examines the content of four sessions: 17th, 58th, 65th, and 100th, providing the history lesson that as the size of the nation grew so did the topics that were reported upon. The 1907 Printing and Binding Act had its own impact; along with other considerations it "authorized by law, the numbering system for the Serial Set worked out by John Ames 40 years earlier," as well as establishing that executive area materials "be designated by title" (DeLong, 130). By the 1980's the *Serial Set* was no longer as important as it had once been and other means of distributing the information made "the *Serial Set* appear to be an anachronism." (DeLong, 133.) DeLong mentions an article by Steven D. Zink (1986) that discusses the difficulties users have in using government documents in general and the *Serial Set* in particular. These difficulties increase the attraction of having catalog records, because the possibility of discovery is greatly increased for those who might never think to select the *U.S. Congressional Serial Set* from their library's database menu. However, there are many libraries that simply cannot afford to purchase an electronic version of the *Serial Set*, so material on use of the print set remain relevant.

2.2 Preservation of the Printed *Serial Set*

Demonstrating that electronic versions have greatly improved access to these government documents are articles that describe the difficulties in using the print version. Thomas, Piscitelli, and Rholes in a 1994 article explore the "Security and Preservation of the *U.S. Congressional Serial Set*." They outline the difficulties patrons have in using the *Serial Set*, and then, after patrons have succeeded in identifying a document, the frustration with discovering that items are missing. Items in the set can be either valuable or fragile and the bindings used on the sets themselves are liable to disintegrate over time. The article gives an overview of concerns with

the accuracy of listings of holdings in depository libraries. Thomas et al. sent a survey in early 1991 with questions about the *Serial Set*. Responses indicated various locations for the set to be housed, with "special collections" and "caged area of documents department" dominating (Thomas, Piscitelli, & Rhodes, 357). Reasons for restriction to access varied. Theft and vandalism were associated with the type of library (regional libraries having the most missing volumes), and age and large size an indicator of an increased likelihood for vandalism (Thomas et al., p. 360). Preservation was more likely to exist in private academic libraries. Thomas et al. provide recommendations for categories in the *Serial Set* to restrict.

An earlier article, from 1988, examined "Use of the *U.S. Serial Set* in an Academic Institution: A Collection Management Tool." Suzanne M. Clark discovered that at her institution, the University of Vermont, the *Serial Set* was only very lightly used and justification for its shelf space was brought into question. More recent volumes (1977-1985) received the most use. Clark advised additional research on use at other institutions, encouraged "use data per population," and cautioned that the intrinsic value of selected volumes be taken under consideration before collection development decisions be finalized.

In an electronic world, access to material in the *Serial Set* is less cumbersome and avoids the peculiarities of the printed set along with the problem of missing or misplaced volumes. Depending upon space and circumstances, libraries that purchase electronic versions of the *Serial Set* may want to maintain selected printed volumes in their collections—securely housed.

2.3 Using the *Serial Set*

An article from 1965 outlines the steps and necessary attention to detail that the strong-hearted researcher, determined to use the microform *American State Papers* or the *Serial Set*, should follow--with assistance from a "softhearted librarian" (Larsgaard, 1965, 304)! The researcher needs to determine a topic and time frame from the very beginning. Details are provided for using the various available print indexes. Somehow, the mind boggles at the 21st century researcher being willing to investigate the 19th century by "attack[ing] logically, patiently and persistently" the multitude of indexes (Larsgaard, 1965, p. 309). Rodney A. Ross (1994) gives recommendations for "Using the *U.S. Congressional Serial Set* for the Study of Western History," informing historians of the American West that its use is essential. Ross's article discusses using the *CIS Index* in order to unearth tidbits from the *Serial Set*. In addition, the *Checklist of United States Public Documents, 1789-1909* is outlined along with other pertinent reference books. Charles D. Bernholz (2009) and Anthony G. Carr describe errors in the *Checklist* regarding Indian Affairs, providing corrections in their article as well as suggestions for finding the resources in the digital Readex *Serial Set*.

The rewards for being logical, patient, and persistent are palpable. There are a number of articles that enumerate the delights in store in the *Serial Set* for the researcher. Jeffery Hartley (2009) covers "Using the *Congressional Serial Set* for Genealogical Research." Land claims, lawsuits, patents granted, pensioners, military registers, and various annual reports provide ample scope for the genealogist to explore. Various land claims are outlined in "Those Elusive Early Americans: Public Lands and Claims in the *American State Papers*, 1789-1837" by Chris Naylor

(2005). These documents provide insight to ancestors who were alive in the years 1789-1837. Difficulties in using the printed *Serial Set* are legion, as enumerated by these researchers; however, they do note the utility of the contents, leading others to be grateful for the library that has digital versions and catalog records to support them.

2.4 Cataloguing the *Serial Set*

After summarizing the contents of the *Serial Set* Aimée C. Quinn (2005a) describes the difficulties with cataloguing government serials by describing the cataloguing of an 1855 document and looks at "digital projects and questions if digitization and the creation of metadata solve the challenges inherent in cataloguing large serial sets" (Quinn, 2005a, 185). She mentions electronic-only 21st century government documents and the challenges presented by them. The number of paper titles decreased to less than half between 1996 and 2003 (p. 190). She then outlines the thought process she had to take to catalog the 1855 item. (The vendor-supplied catalog records that CSUL purchased are looking very attractive at this point.) Quinn then discusses the various ongoing digitization projects and any accompanying metadata, including that done by Readex. This is the only article which discussed cataloging of electronic government documents in detail. More on this topic will be useful additions to the field.

"Knowing Where They Went: Six Years of Online Access Statistics via the Online Catalog for Federal Information" discusses click-throughs from a library catalog that provided URLs for government documents that exist on the Web. The provision of incorrect URLs were a problem part of the time, but the author concludes that even with these setbacks users preferred digital content over print, so adding additional resources beyond readily available GPO offerings were

useful and used. The library also loaded the first half of the Readex *Serial Set* catalog records, but the article does not mention massaging them. Click-throughs from their online catalog demonstrated that their *Serial Set* and other collections are being accessed through their library catalog (Brown, 2011).

2.5 Notices and Reviews of the Online *Serial Set*

This section will be helpful for librarians that are considering the purchase of the electronic *Serial Set* from Readex for their libraries. A brief notice of the forthcoming digital *U.S. Congressional Serial Set* from Readex was published in the September 2003 *Computers in Libraries*. In Spring 2004, mention of digitizing the *Serial Set* (Readex and LexisNexis) was made in *DttP: Documents to the People* (Rossman, 2004). The *Serial Set* as a part of Readex's "Archive of Americana" was also mentioned in the Product News & Reviews section of the December 2004, *Information Today*. Two separate news briefs outlined the information that Readex was "digitizing from the original print volumes . . . housed in Dartmouth's Baker-Berry Library" and "not from microform editions" (Association of College & Research Libraries, 2006; Millwood Group Corp., 2006).

An examination of the Readex *Serial Set* was paired with a review of the LexisNexis version in a Summer 2005 review in *DttP: Documents to the People* (Miller, 2005). *Government Information Quarterly* published an eleven page in-depth introduction and then a separate review of both digital versions of the *Serial Set* in 2005. The reviewers were asked to consider ten questions in their reviews of the databases (Quinn, 2005b; Clausen, 2005; Meister, 2005). *The Charleston Advisor* also weighed in with a double review for the Readex and LexisNexis

versions of the *Serial Set* written by Jian Anna Xiong (2004). The interface, search features, search results, and added features were all addressed. *Choice* provided a short summary of features in its reviews, one for the *U.S. Congressional Set* (Stuart, 2010) and the other for the *American State Papers* (Hardenbrook, 2006). In *Microform & Imaging Review* Barbara Norelli (2005) provided a glowing and detailed review of the digital Readex *Serial Set*; she covered options for searching, display of results, searching within results, printing and downloading options, usability, etc. In Sarah Santos's (2005) review, she covered facts on the Readex and LexisNexis products, providing information on Readex's "added value to the collection by identifying and correcting errors, as well as identifying gaps that occurred" and "human indexing and examination of each publication."

2.6 E-Book Cataloging

While not directly related to the *Serial Set*, e-books have had a growing presence in libraries, and the cataloging of them has created a new set of concerns for librarians. An editorial in *Library Collections, Acquisitions, & Technical Services* described the University of Chicago Library's thought process on vendor-supplied catalog records. Getting materials processed quickly, getting their records in the catalog, and adding these records to nationally available resources, were priorities (Mouw, 2005). A pilot project that involved the purchase of vendor records for materials in Italian did not allow the sharing of records, a grave concern to Mouw. Two articles, both published in *Library Resources & Technical Services* discussed the management and challenges of using vendor records for e-books. Annie Wu (2010) and Anne M. Mitchell provided a literature review which included items on batch processing, using AACR2, and more. They discussed what they call "provider-neutral" (164) e-book records and guidelines, such as

dates for e-books being "based on the original monograph, whether print or electronic" (168). Wu and Mitchell's article discussed batch processing in detail. Kristin E. Martin (2010) and Kavita Mundle wrote on the challenges of managing MARC e-book records from Springer. Sanchez et al. (2006) discussed cleaning up NetLibrary cataloging records. Problems and solutions were described in detail. There are some, but not many other articles that focused on specific cases; for example, loading records for freely available digital books (Beall, 2009). They are not included because this article is focusing on government documents, which have specific kinds of challenges; therefore other types of cataloguing problems, such as reconciling publication and digitization dates, a greater problem in the commercial vendor setting, are not relevant here.

2.7 Batchloading Records and European Vendor Records

Mary Finn (2009) described batch loading as contributing to authority control management problems; she provides specific examples and shares Virginia Tech's steps for correcting problems. Batchloading electronic and microform records was the topic of Rebecca L. Mugridge (2009) and Jeff Edmunds's article. Sources for MARC records were described along with the desire from subject specialists to have bibliographic record sets added to the library catalog. The authors provided details on record quality, format duplication, purchasing vs. downloading, related decision making, workflow, and more.

Charlene Kellsey (2002) examined vendor provided records for foreign-language materials found in OCLC. The number of records, their classification (LC call number), and the specific languages were tallied. Kellsey (2001) first examined Italian monograph records in OCLC.

Batch loading has its own challenges as the final sections of this article will describe. This literature review has shown how the long and historically relevant content of the *Serial Set* makes the electronic version a prime candidate for inclusion in a library's online catalog. As discovered at CSUL, once the records were loaded, a significant increase in the number of documents viewed was evident.

3. Readex-provided MARC Records

Readex's *U.S. Congressional Serial Set* collection is supplemented by 363,238 MARC records (A. A. Imholtz, Jr. personal communication, June 21, 2011). These records cover all publications between 1817 (15th Congress) and 1980 (96th Congress). The records were made available in early fall 2007 and purchased by CSUL late that same year. In 2008, CSUL loaded bibliographic records through 1919 (65th Congress). This group consists of 194,725 bibliographic records; these are subsequently divided into eight sections. Each of these sections varies in size from approximately 17,000 to 31,000 records. In addition, Readex sells, separately, 6,278 MARC records for the *American State Papers* (A. A. Imholtz, Jr. personal communication, June 21, 2011). These were purchased by CSUL soon after they were made available by the vendor in late summer 2009.

Although the acquisition of the MARC records provides numerous benefits, as seen in section 4, there are many difficulties involved ingesting the records into the library's system. Initially, discussions were held between the Libraries' Metadata Librarians regarding a sample of the records and logistical concerns of loading such a massive collection. In analyzing the sample in

a text editor, other problems were discovered. It was decided that some of the problems would best be addressed prior to loading the records, whereas authority control problems would have to wait until after loading. CSUL determined that it was worth it to make improvements to the Readex records because the enhancements impact patron discovery and reinforce the integrity of the catalog.

System limitations must be addressed when batch loading large record sets, especially when they include hundreds of thousands of records. Innovative Interfaces Incorporated Millennium (III) works on a transaction file that records all activities within the system until a full backup is run which clears the file. Once the file reaches 100 percent the entire Integrated Library System (ILS) does not allow anymore transactions and stops working. At CSUL, full backups are run every week day. For the *Serial Set*, 10,000 records were test loaded to see the impact on the transaction file. This test added 25 percent to the transaction file, so the librarian primarily responsible for the loading of the Readex records and the system administrator agreed to daily loads of approximately 20,000 records. This would allow all other library activities to take place and allow flexibility in splitting the batches provided from Readex. The initial eight files from Readex containing 194,725 MARC records were divided in 14 batches and loaded over a three week time period.

The following is an example of a Readex supplied MARC record:

```
=LDR  02830cam  22005291i 4500
=001  NB00000144193
=003  Readex
=005  20071001171026.3
=006  m\\\\\\\\d
=007  cr\cn\||||||||
=008  070221s1817\\\\dcu\\\\\\s\\\f000\0\eng\d
=040  \\$aReadex$cReadex
=110  1\$aUnited States.$bCongress.$bSenate.
=245  10$aReport on the Japanese naval medical and sanitary features of the Russo-Japanese War to the
```

Surgeon General, U.S. Navy, by Surgeon William C. Braisted, U.S. Navy. Tokyo. July 15, 1905.$h[electronic resource]
=260 \\$aWashington, DC,$c1905
=300 \\$a82 p. :$billustrations, maps, tables.
=440 \0$aUnited States congressional serial set ;$vserial set no. 4915
=490 1\$aSenate document / 59th Congress, 1st session. Senate ;$vno. 328
=500 \\$aTable of contents, p. 3.
=533 \\$aElectronic reproduction. $bChester, Vt.: $cNewsBank, inc., $d2005. $nAvailable via the World Wide Web. $nAccess restricted to Readex U.S. Congressional Serial Set subscribers.
=540 \\$aCopyright 2007 by NewsBank, Inc. All rights reserved.
=610 27$aNavy Medical College at Tokyo, Japan.$2Readex congressional thesaurus
=650 07$aCommunicable diseases.$2Readex congressional thesaurus
=650 07$aHospital ships.$2Readex congressional thesaurus
=650 07$aHospitals, Naval and marine.$2Readex congressional thesaurus
=650 07$aHospitals.$2Readex congressional thesaurus
=650 07$aMedical care.$2Readex congressional thesaurus
=650 07$aMedical education.$2Readex congressional thesaurus
=650 07$aMedical instruments and apparatus.$2Readex congressional thesaurus
=650 07$aMedicine, Military.$2Readex congressional thesaurus
=650 07$aNavies.$2Readex congressional thesaurus
=650 07$aQuarantine.$2Readex congressional thesaurus
=650 07$aSanitation.$2Readex congressional thesaurus
=650 07$aSurgery.$2Readex congressional thesaurus
=650 07$aTransport of sick and wounded.$2Readex congressional thesaurus
=650 07$aWar casualties.$2Readex congressional thesaurus
=650 07$aWarships.$2Readex congressional thesaurus
=650 07$aRusso-Japanese War (1904-1905)$2Readex congressional thesaurus
=650 07$aJapanese (People)$2Readex congressional thesaurus
=650 07$aNurses.$2Readex congressional thesaurus
=651 \7$aJapan.$2Readex congressional thesaurus
=655 \7$aExecutive Department Publications.$2Readex congressional thesaurus
=655 \7$aMonographs.$2Readex congressional thesaurus
=700 1\$aBraisted, William C. (William Clarence)
=710 1\$aUnited States.$bNavy.
=830 \0$aSenate document (United States. Congress. Senate) ;$v59th Congress, no. 328
=856 40$uhttp://docs.newsbank.com/select/serialset/1169A3B93069B510.html

4. Readex MARC Record Pluses

The MARC records provided by Readex include many positive traits. Previously, effectively finding *Serial Set* materials via a library catalog was nearly impossible without first consulting the 52 volume *Congressional Information Service (CIS) United States Serial Set Index*. The MARC record set drastically alters this experience in allowing access never before possible. All the records are cataloged with a high level of detail, consistency, and thoroughness. Numerous subject headings and added authors are provided on all records. Many of the records are even longer than the text they describe. The records are formatted consistently with standardized fixed fields, series, and electronic resource description.

5. Readex MARC Record Problems

CSUL implemented analyzing all batch loads prior to loading them into the library's ILS. First, a Readex *Serial Set* sample was uploaded into a text editor for analysis. When several drawbacks were discovered, they were divided into two groups relating to patron access impact. The low impact group consisted of relatively inconsequential errors, for example antiquated formatting, non-ISBD (International Standard Bibliographic Description) punctuation, and words not abbreviated per AACR2 (*Anglo-American Cataloging Rules*, 2nd edition). Low impact problems were not fixed. The high impact group related to problems that would alter patron access. These problems consisted of numerous encoding errors and hundreds of thousands of unauthorized headings. The high impact error group was additionally analyzed for the feasibility of fixing the problems. For instance, time and staff demands were addressed concerning altering the substantial authority record inaccuracies. Each high impact problem is summarized below with the accompanying CSUL solution. Readex was not consulted. It is important to note that CSUL's Integrated Library System (ILS) is Innovative Interfaces Incorporated's Millennium. Therefore, the problems below could differ for libraries using other systems. The format used below to describe the workarounds done at CSUL is first, an outline of the (high impact) problem, and then the solution, phrased as what to do to solve the problem. Other libraries may choose to do only a few of these fixes; CSUL thought it worthwhile to make these corrections.

5.1 Multiple title (245) fields

- **Problem:** Readex enhanced older monographic records to create electronic resource records. When the title was erroneously broken into multiple 245 fields, since many

eighteenth and nineteenth century titles contained exhaustive detail, they needed to be corrected. For example: *In the Senate of the United States. March 29, 1852. Ordered to be printed. Mr. Mason made the following report: (To accompany Bill S. No. 323.) The Committee on Foreign Relations, to whom was referred the resolution of the Senate, instructing them "to inquire into the propriety and justice of providing by law, pursuant to the recommendation of former presidents of the United States, and last by President Polk, in his message of the 7th December, 1847, for the payment of the claim there mentioned as arising to certain Spanish subjects, in the case of the schooner 'Amistad,'" have had the same under consideration, and submit the following report...* When these electronic resource records were corrected, a GMD (General Material Designation) subfield h was only created in the first title (245) field.

- **Solution**: After loading, create a list of the loaded Readex records. Run a second list against the first list's results, and search for title (245) fields that do not include the subfield h. This provides a list of records with multiple title (245) fields. Since we were unable to devise an automated way to correct the records, they were manually fixed by cutting each second title (245) field and pasting it to the end of the first title field.

5.2 Title field (245) second indicator zero with titles beginning with a, an, or the

- **Problem:** Approximately three thousand records begin their titles with one of the articles "a," "an," or "the" without a skip value. This means that none of these records are accessible when searching by title. This is a problem for researchers who have a citation

with the title (and pretty much all of them will) who search the library catalog for the item by title (which tends to be the fastest search in online library catalogs).

- **Solution:** The most straight forward method we devised to fix this problem is after loading the records. Create record lists with each article and a succeeding space plus a line to only harvest records from this set. After completing the list, globally change the second indicator and fixed field skip value.

5.3 Dollar sign used in text

- **Problem:** The dollar sign ($) is used as a delimiter by III's input ports. As a consequence, the eleven *Serial Set* records that contain dollar amounts in the records were mis-loaded and these dollar signs would become delimiters after loading the records.
- **Solution:** In MarcEdit, generate a field count for the record set and analyze the numbered subfields. Twelve instances were discovered in the 500 and 630 fields. A load table was created to retain the dollar signs and specifically apply to the above two fields.

5.4 Fixed field date

- **Problem:** Readex placed the date 1817 in all the first date fields of each record's 008. Since CSUL uses the 008 date fields in the patron interface facets, this would create the erroneous information that over 194,000 publications were published in 1817.
- **Solution:** There was not an ideal automated way to correct the problem, so the decision was made to change the dates after loading the records into the library's catalog. Each

record contained the correct date in the 260 subfield c. This made it possible to create a list for year and apply a global update to fix the problem.

5.5 Languages

The multilingual documents contained in the *Serial Set* and *American State Papers* help exemplify the breadth of materials encompassed in these publications. Multilingual sources include treaties, agreements, correspondence, reports, and publications containing specific Native American languages. Generally, the text is either presented first followed by the English translation (see Figure 1 and 2), or the text and translation are given in a parallel format (see Figures 3-5). *New Mexico: Letter from the Secretary of the Territory of New Mexico, transmitting copies of the acts, resolutions, &c., of that territory. January 9, 1852* (Figures 1 and 2) show acts written in Spanish followed by English translation. *Treaty with the King of the Belgians* (Figure 3) from 1865 illustrates the use of parallel text in the *Serial Set*. Figures 4 and 5 demonstrate how the extensive Bureau of American Ethnology bulletins combined formats. *Alsea texts and myths* (Frachtenberg, 1918) show a common way Native American texts are displayed. The original text or phonetic transcription of Alsea is specifically translated underneath each word in English and the entire passage is followed by a full narrative translation.

Figure 1
Page 1 of New Mexico: Letter from the Secretary of the Territory of New Mexico, transmitting copies of the acts, resolutions, &c., of that territory. January 9, 1852. Readex Record Number: 104925B24BE07B48

ACTS, RESOLUTIONS, AND MEMORIALS

OF

THE LEGISLATIVE ASSEMBLY OF THE TERRITORY OF NEW MEXICO,

Passed at a session begun and held June 2, A. D. 1851.

JOINT RESOLUTION.

Whereas, by resolutions of the Council and House of Representatives of the Territory of New Mexico, it appears that the legislative assembly is favorable to contracting a loan of ten thousand dollars, to defray the expenses of the legislature and for other purposes; therefore,

Be it resolved, That a joint committee of two members from each house be empowered and authorized to contract for said loan of ten thousand dollars, giving bonds for the same, payable at six months after date, or sooner if the treasurer shall be able so to do, which said bonds shall be signed by the treasurer of this Territory and countersigned by the governor, and shall pledge the appropriation as already made by the United States for the payment of the same; and further, that the said committee shall pay over said sum of money to the treasurer of the Territory when procured; and further, that the auditor of public accounts be required to audit, and the treasurer to pay, out of the sum of money aforesaid, such accounts as have been necessary for the current expenses of the legislative assembly.

THEODORE D. WHEATON,
Speaker of the House of Representatives.
JOHN F. ORTIZ,
President of the Council.

Approved June 17, 1851:
JAMES S. CALHOUN, *Governor.*

UN PROYECTO que se titulará un decreto estableciendo el empleo de traductor, para el Territorio de Nuevo Méjico.

Decrétese por la Asamblea Legislativa del Territorio de Nuevo Méjico:

SECCION 1. Que será nombrado por el gobernador, con el consentimiento i aprobacion del concejo legislativo, un traductor para dicho territorio, capaz de traducir, del idioma Castellano al Ingles, i vice versa.

SECCION 2. Será el deber de dicho traductor, suplir traducciones propias de todo proyecto, resolucion, ó documento, perteneciente á los

Figure 2
Page 2 of New Mexico: Letter from the Secretary of the Territory of New Mexico, transmitting copies of the acts, resolutions, &c., of that territory. January 9, 1852. Readex Record Number: 104925B24BE07B48

negocios de la asamblea legislativa, cuando se le pida. Tambien, dirijir la imprenta ó publicacion de la parte Castellana de las leyes pasadas por, i los diarios de dicha asamblea legislativa.

SECCION 3. Será el deber de dicho traductor, suplir al departamento del ejecutivo, i al secretario del Territorio, con traducciones de cualesquier papel ó documento pertenecientes á dichas oficinas, cuando se le pidan.

SECCION 4. Dicho traductor recibirá por sus servicios, la suma de dos mil pesos, anuales pagaderos por trimestre, i que dicha suma sea i por estas presentes está appropriada de los fondos para gastos incidentales de la asamblea legislativa, como ha proveido el Congreso de los Estados Unidos.

SECCION 5. A dicho traductor, le será permitido un asistente, que recibirá, tres pesos diarios, por sus servicios, i será pagado de la misma manera, i de los mismos fondos, como se ha proveido para el traductor, pero será ocupado en solo los negocios de la legislatura, i no en otros, no excediendo de tres meses en el año.

SECCION 6. Ynmediatamente que se apruebe la presente lei, se procederá al nombramien to de traductor.

THEODORE D. WHEATON,
Speaker of the House of Representatives.
JUAN F. ORTIZ,
Presidente del Consejo.

A BILL to be entitled "An act establishing the office of translator for the Territory of New Mexico."

Be it enacted by the Legislative Assembly of the Territory of New Mexico:

SECTION 1. That there shall be appointed by the governor, with the advice and consent of the legislative council, a translator for said territory, capable of translating from the Spanish into the English languages, and *vice versa.*

SEC. 2. It shall be the duty of said translator to furnish correct translations of all bills, resolutions, or documents pertaining to the business of the legislative assembly, when called upon so to do; also, superintend the printing or publication of the Spanish portion of the laws passed by, and the journals of, said legislative assembly.

SEC. 3. It shall be the duty of said translator to furnish the executive department and the secretary of the Territory with translations of any paper or document pertaining to said offices, when required so to do.

SEC. 4. Said translator shall receive for his services the sum of two thousand dollars annually, payable quarterly, to be, and is hereby, appropriated from the funds appropriated by the Congress of the United States for contingent expenses of the legislative assembly.

SEC. 5. Said translator shall be allowed an assistant, who shall receive three dollars per diem for his services, to be paid in the same manner, and from the same funds, as provided for the translator; but

Figure 3

Treaty with the King of the Belgians. Message from the President of the United States, transmitting a copy of two treaties between the United States and the King of the Belgians in relation to the Scheldt dues, &c. January 9, 1865. Readex Record Number: 108B1AAB87997EE0

The President of the United States of America, on the one side, his Majesty the King of the Belgians, on the other side, having deemed it advantageous to complete, by new stipulations, the treaty of commerce and navigation entered into by the United States and Belgium on the 17th day of July, eighteen hundred and fifty-eight, have resolved to make a convention in addition to that arrangement, and have appointed for their plenipotentiaries, namely:

The President of the United States, Henry Shelton Sanford, a citizen of the United States, their minister resident near his Majesty the King of the Belgians; his Majesty the King of the Belgians, the Sieur Charles Rogier, grand officer of the Order of Leopold, decorated with the iron cross, grand cross of the order of the Ernestine Branch of Saxony, of the Polar Star of St. Maurice and St. Lazarus, of Our Lady of the Conception of Villa-Vicosa, of the Legion of Honor, of the White Eagle, &c., a member of the Chamber of Representatives, his minister of foreign affairs; who, after having communicated to each other their full powers, found to be in good and proper form, have agreed upon the following articles:

ARTICLE I.

From and after the day when the capitalization of the duties levied upon navigation in the Scheldt shall have been secured by a general arrangement—

1st. The tonnage dues levied in Belgian ports shall cease;

2d. Fees for pilotage in Belgian ports and in the Scheldt, in so far as it depends on Belgium, shall be reduced twenty per centum for sailing vessels, twenty-five per centum for vessels in tow, thirty per centum for steam vessels;

3d. Port dues and other charges levied by the city of Antwerp shall be throughout reduced.

ARTICLE II.

In derogation to the ninth article of the treaty of the seventeenth of July,

Le Président des Etats-Unis d'Amérique, d'une part, et sa Majesté le Roi des Belges, d'autre part, ayant jugé utile de compléter par de nouvelles stipulations le traité de commerce et de navigation signé entre les Etats-Unis et la Belgique le dix-sept Juillet mil huit cent cinquante huit, ont résolu de conclure une convention additionnelle à cet arrangement et ont nommé pour leurs plénipotentiaires, savoir:

Le Président des Etats-Unis, Henri Shelton Sanford, citoyen des Etats-Unis, son Ministre Résident près sa Majesté le Roi des Belges; sa Majesté le Roi des Belges, le Sieur Charles Rogier, Grand Officier de l'Ordre de Léopold, décoré de la Croix de Fer, Grand Cordon des Ordres de la Branche Ernestine de Saxe, de l'Etoile Polaire, des SS. Maurice et Lazare, de Notre Dame de la Conception de Villa-Viçosa, de la Légion d'Honneur et de l'Aigle Blanc, &a., Membre de la Chambre des Représentants, son Ministre des Affaires Etrangères, lesquels, après s'être communiqué leurs pleins pouvoirs, trouvés en bonne et due forme, sont convenus des articles suivants:

ARTICLE Ier.

A partir du jour où la capitalisation du péage de l'Escaut sera assurée par un arrangement général:

1°. Le droit de tonnage prélevé dans les ports belges cessera d'être perçu;

2°. Les droits de pilotage dans les ports belges et dans l'Escaut, en tant qu'il dépendra de la Belgique, seront réduits; de 20 p. c. pour les navires à voiles, de 25 p. c. pour les navires remorqués de 30 p. c. pour les navires à vapeur;

3°. Le régime des taxes locales imposées par la ville d'Anvers sera dégrévé dans son ensemble.

ARTICLE II.

Par dérogation à l'art: 9 du traité du dix-sept Juillet mil huit cent cin-

Figure 4
Page 22 of Leo J. Frachtenberg's Alsea texts and myths. Readex Record Number: 11926DD88A647620

CREATION MYTHS

1. YEHAᵉMĪYŪ [1]

CYCLONE

(Told by Tom Jackson in 1910)

Xa'mEt-s-mEsha'lslatsLō yā'tsx. Sūdāᵃ'st ts-ła'mxadōok·.
One it is a woman old is living. Five her children her.

Qauwāᵉᵃ qā'altsuxs.[2] Xam taiᵉ mukwaᵉslī. MEā'łtxudau'xaLx.
All male (plural). One only woman. Continually big keep on they.

Tem mi'siLx LEā'łtxudī'yū, temi'Lx pxēltsūsaī'nx ats-Lī'ak·iLx.
Finally when they big became, then they ask begin her that their mother their.

"K·Eł x·ūlamtxaī'm hau'k·s ⁱk·tas leʻwī'."—"K·ip-E'n tā'xtī,"
"Will we travel begin will everywhere to over this world."— "Will you what do,"

ts-yEaī'sk· as mukwaᵉslītsLō. "Ā'a, ł-qauwāᵉᵃ hī'k·e ta'xtī."—
her speech her the woman old. "Yes, we every just do."—

"Sin haⁿ'yauk·,[3] k·Ep ⁱLīyaᵉ hī'k·e qauwāᵉᵃ ta'xtī hilkwaīsa'a,
"My mind inside, will you not just every thing do will it,

sips [4] ayaī'mi. K·ip-E'n mūⁿ'łiī ta'xtī?", ts-yEaī'sk· as
if you go will. Will you what by way of imagining do?", her speech her that

mEsha'lslatsLō. —"Ā'a, k·Eł łpū'pEnhaut. K·Ełts-axa [5]
woman old.— "Yes, will we shinny playing. Habitually we again

LEyaⁿ'hats!īt; k·Ełts-axa ītaᵉ LEmk·i'łhīxamt." [6]—"K·eai'sa,
guessing; habitually we again also dancing." —"All right,

p-ayaī'mi. P-spaī'dī ku'pin qā'sint." Temi'Lx muⁿ'hū
you go shall. You take her along will this your younger sister." Then they finally

k·e'a ī'mstE łt!a'msīyū. Temuⁿ'hū [7] yEaī'nx xas mEsha'lslatsLō
indeed thus agree quickly. And now tells her the woman old

ats-pīya'tsEk·. "K·iLx [8] ūsta'a kuha'm hā't!ōo. K·iLx [8]
that her daughter her. "Shalt them follow (fut.) these thy elder brothers. Shalt them

ts!ōwaī'Lī-slō [9] xe'īLk·e, k·iLx ⁱLīyaᵉ nā'k·eai qaⁿ'yEkwai!
watch (fut.) all over carefully, shall them not anywhere harm them!

Tū! spaī'tx tas ti'nExtsū! K·ip qō'tsE sā'lsxumī."
Here! take along this cane! Will you it (as) magic power use it (fut.)."

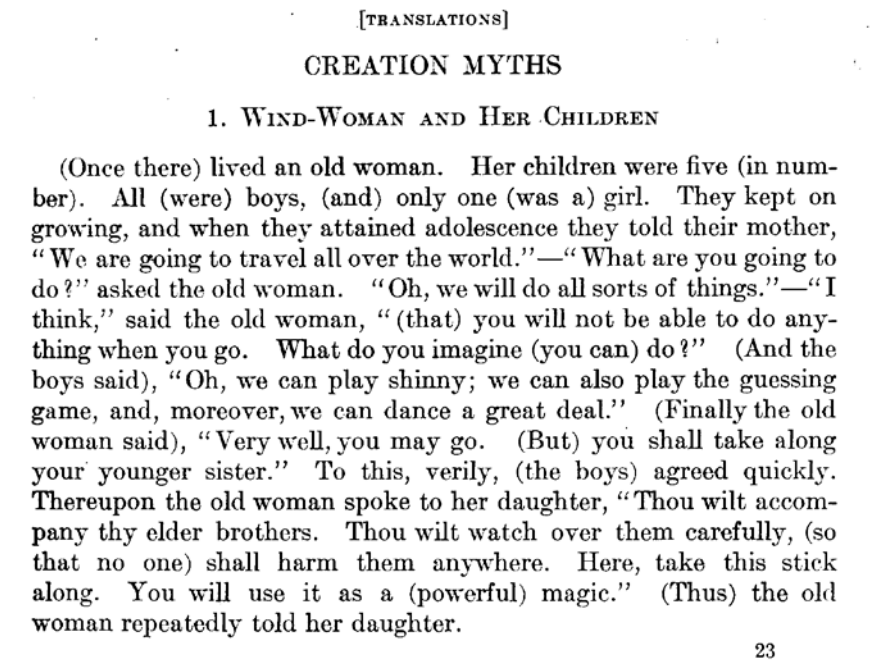
TsimyEaī'sk· [10] ats-pīya'tsEk· xas mEsha'lslatsLō.
Her continual speech her this her daughter her the woman old.

---

[1] Compare Boas, Traditions of the Tillamook Indians, p. 138; Sagen, pp. 1, 16, 19, 47, 56, 63.
[2] *qā'alt* MALE BEING; *-tsuxs* collective.
[3] *hainᵉ* MIND; *-auk·* suffixed particle INSIDE.
[4] *sis* conditional particle; *-p* 2d person plural.
[5] *k·Ets* temporal particle; *-ł* first person plural; *-aza* suffixed particle.
[6] *L-* prefix; *m-* prefix; *k·īlh-* TO DANCE; *-ī* inchoative; *-xamt* infinitive.
[7] Contracted for *tem muⁿ'hū*.
[8] Instead of *k·ExiLt*.
[9] Stem *ts!ōu- . . . -slō* TO WATCH, TO NOTICE.
[10] *ts-* possessive; *m-* continuative: *yEa-* TO SAY; *-aī* inchoative; *-s* nominal; *-k·* possessive.

22

Figure 5
Page 23 of Leo J. Frachtenberg's Alsea texts and myths. Readex Record Number: 11926DD88A647620



[TRANSLATIONS]

CREATION MYTHS

1. WIND-WOMAN AND HER CHILDREN

(Once there) lived an old woman. Her children were five (in number). All (were) boys, (and) only one (was a) girl. They kept on growing, and when they attained adolescence they told their mother, "We are going to travel all over the world."—"What are you going to do?" asked the old woman. "Oh, we will do all sorts of things."—"I think," said the old woman, "(that) you will not be able to do anything when you go. What do you imagine (you can) do?" (And the boys said), "Oh, we can play shinny; we can also play the guessing game, and, moreover, we can dance a great deal." (Finally the old woman said), "Very well, you may go. (But) you shall take along your younger sister." To this, verily, (the boys) agreed quickly. Thereupon the old woman spoke to her daughter, "Thou wilt accompany thy elder brothers. Thou wilt watch over them carefully, (so that no one) shall harm them anywhere. Here, take this stick along. You will use it as a (powerful) magic." (Thus) the old woman repeatedly told her daughter.

23

- **Problem:** A few thousand records include languages other than English. Readex used 546 fields in describing multiple languages; however, the comparable codes were not added to 041 fields. Although this does not impact the traditional online public access catalog (OPAC), it affects facet searching in the library's VuFind discovery tool. This problem holds significance in that these multi-language documents generally contain primary sources, such as diplomatic relations and agreements. The documents are often provided in parallel text, making the comparison between versions easy to distinguish. Since this is primary source material, it is important to indicate the languages used in the documents for patrons.
- **Solution:** No automated way was conceived to add the 041 language codes; therefore, the problem was addressed after loading the records. A list was created by searching for 546

fields in the *American State Papers* and *Serial Set* and approximately 1000 records were identified. The records were manually edited, adding 041 fields to each record, making the records capable of being retrieved through a search by language, thereby improving access to primary materials.

5.6 Subject Headings

Readex uses a Readex Congressional Thesaurus for all subjects. The 6xx fields are formatted as follows: 650 07 $a Treaties $2 Readex congressional thesaurus. The initial batch of *Serial Set* records were loaded without altering these subjects, since few matched Library of Congress Subject Headings (LCSH), and it was seen as too laborious to successfully change hundreds of thousands of headings to LCSH. Additionally, many of the personal and corporate author headings were inconsistent for authority control. The decision was positive from a database management standpoint; however, CSUL's Millennium system does not allow searches for 6xx fields that are not LCSH. Although the Readex subjects are indexed as keywords in the OPAC, they are difficult to harvest with the staff interface. Any important subjects to CSUL patrons are difficult to change and make consistent with the catalog's existing subjects. Since CSUL had tens of thousands of *American State Papers* and *Serial Set* records to load, it was decided to reexamine the subjects, and see if it would be beneficial to load the records as LCSH.

After receiving the *American State Papers* record set, a test was conducted on the subject headings. One hundred headings were taken from twenty random records and evaluated. The results were categorized in three groups, headings that matched LCSH or name authority

records; “see reference” headings from authority records; and headings that did not match authority records.

Table 1
Analysis of American State Papers Readex Thesaurus Subject Headings

MARC Field	Correct Headings	See References	Incorrect Headings
600	3	0	7
610	7	0	4
630	0	0	5
650	18	13	26
651	8	0	9
Totals	36	13	51

Personal and corporate name subjects were consistent with the problems found in analyzing author authorities, see section 5.9. The headings used by Readex for uniform titles (630 field) are consistently formatted inconsistently to Program for Cooperative Cataloging (PCC) standards. For instance, Readex lists all treaties under the 630 field, e.g. 630 07 $a Treaty with the Shawnee. $d 1786; however, PCC constructs treaties as corporate headings, e.g. 610 20 $a Shawnee Nation. $t Treaties, etc. $g United States, $d 1786 Jan. 31. These are often difficult to find and correct, since the Readex headings do not match references in the authority records.

The Readex subjects in 650 and 651 fields pose several problems if they are converted to LCSH. Although nearly one-third of the headings matched LCSH, these headings are often words providing little description and no qualifying information, e.g. Claims, Census, Explorers, or Fortifications. This is also the case for the “see reference” headings. Nearly half of the subjects are incorrect and do not provide a straightforward way to correlate to LCSH. The 651 fields provide geographical headings that would generally serve patrons

better if they were attached to the above 650 fields. Instead, the general headings for countries, states, and provinces have little benefit for a patron. The more granular headings are often formatted differently from LCSH. Cities and rivers are qualified with the full state name(s), e.g. Portland, Maine compared to the LCSH where the state abbreviation is used, e.g. Portland (Me.). Readex uses territory names and dates (e.g. Mississippi Territory (1798-1817)), whereas in LCSH, territories are generally listed as "see" references under state name, or when they are used they are constructed without dates impacting indexing and results for patrons.

After reviewing the sample results of the text run, it was decided to continue loading uniform titles and 650/651 fields as Readex Thesaurus headings because they would be too time consuming to change. However, name authorities would be changed to LCSH. The uniform titles were consistently incorrect and offered little or no possibility to be easily converted to authorized headings. Similarly, the 650/651 headings were seldom correct and were deemed too laborious to correct after changing to LCSH. In contrast, selected personal and corporate names were altered to provide greater patron access. The subject headings for prominent figures were modified to match authorized LC headings. Along with presidents, generals, and other prominent historical figures, persons with local or regional importance were given special consideration. These problems were already being addressed through their use in author fields; therefore, correcting the subject authors would simply be coupled to the existing author global corrections.

5.7 Creating proxy URLs

The records include 856 links to each resource; however, the links need to be amended with a local proxy or WAM (Web Access Management) URL for granting authorized patrons access offsite. This can be accomplished by creating a load table to supplement the load profile. The automated process saves time and prevents possible errors in making the changes in a text editor or post-load global updating.

5.8 Classification

Readex did not put classification numbers in either the *Serial Set* or *American State Papers*. The decision was made to add the general SUDOC classification Y 1.1/2: to each record and the load profile performed the task. This enables the possibility of identifying all of the records by call number in the library catalog. This also serves as a discovery replacement for the de-accessioned print *Serial Set*.

5.9 Author Authority Control

The Readex records have hundreds of thousands of unauthorized author headings. These include both personal and corporate authorship in 1xx and 7xx fields. Unauthorized records means finding *Serial Set* and non-*Serial Set* items under one author is not possible (they are not collocated). A brief analysis was conducted to illustrate the extent of these authority problems. This will be illustrated by three examples: 1) personal entry, 2) corporate entry, and 3) presidential entries.

5.9.1 Personal Entry: Grant, Ulysses S.

The Readex record set includes two headings for the above name: Grant, Ulysses S., $d 1822-1885 and Grant, Ulysses S. These headings illustrate two key problems with the Readex records' personal authority entries.

- **Problem 1:** Although the first entry includes qualifying dates, the authorized form of the heading is not used. Additionally, the form used is not included in the authority record's cross references. As a result, this incorrect heading would remain undetected by any automated process.
- **Problem 2:** The Readex MARC records include countless personal author entries similar to the second Grant heading which excludes any qualifying information. This example is particularly confusing since there are six Ulysses S. Grants with authority records born in the 19th century. Fortunately, all the works relate to geological data published in the first decade of the 20th century which points to Grant, U. S. $q (Ulysses Sherman), $d 1867-1932.
- **Solution:** Locate and differentiate all U.S. Grant headings. Once the separate entries are grouped together, globally change each group to the correct authorized heading.

5.9.2 Corporate Entry: United States. War Department.

In reviewing all the United States War Department headings, the major problems relate to using authorized headings. There are approximately 7,000 War Dept. entries that use an unauthorized form of a corporate author. These are all cross-referenced and most should be fixed by the library's automated authority service. An exception will be the headings

with incorrect or unrecognized qualifying information, including dates (Problem 1) and subordinate units (Problem 2).

- **Problem 1:** A common practice throughout the Readex MARC records is qualifying corporate authors with dates. This practice is not only improper formatting for corporate authors, but it inhibits the ILS from determining the incorrect heading from the authority record's cross-reference. There are two ways Readex appended dates to corporate author headings. First, a $g (miscellaneous information) is added to 110 fields. Second, date ranges are added to 710 $b after the subordinate unit.

  Example 1

  - Incorrect heading: United States. $b Congress. $b Senate. $b Committee on Military Affairs. $g (1816-1857; 1868-1946).
  - Correct heading: United States. $b Congress. $b Senate. $b Committee on Military Affairs.

  Example 2

  - Incorrect heading: United States. $b War Dept. $b Office of Indian Affairs (1824-1849)
  - Correct heading: United States. $b Office of Indian Affairs.

- **Solution:** For example 1, search for the $g in the Readex records and globally delete the $g. For example 2, use the following Millennium search query to find all government agencies with $b dates, and this should exclude all corporate presidential entries: BIBLIOGRAPHIC MARC Tag 110|b matches [0-9)]{5} OR BIBLIOGRAPHIC MARC Tag 710|b matches [0-9)]{5}

Figure 6

Screenshot from Innovative Interface's Inc. Millennium ILS in Create Lists

Type	Field	Condition	Value A
BIBLIOGRAPHIC	MARC Tag 110|b	matches	[0-9)]{5}
BIBLIOGRAPHIC	MARC Tag 710|b	matches	[0-9)]{5}

- **Problem 2:** The records include numerous corporate entries with incorrect subordinate units.
  - Incorrect heading: United States. $b War Dept. $b Office of the Chief Signal Officer. $b Bureau of Aircraft Production.
  - Correct heading: United States. $bArmy. $b Signal Corps. $b Bureau of Aircraft Production
- **Solution:** These problems are often difficult to find, since the entry is not cross referenced in the heading's authority record. Many headings use multiple incorrect subordinate units and one effective way to find the correct heading is to work from the lowest unit and work backward. Generally, the lowest given unit is correct, although it is placed under the wrong command.

5.9.3 Corporate Entry: United States. Navy Department.

Many corporate headings used in the *Serial Set* records include the contemporary heading for government agencies. The Navy Department exemplifies this problem. Although "United States. Navy Dept." should be used on all related materials prior to 1948, Readex exclusively used only the succeeding heading "United States. Dept. of the Navy." This error means that the entries are not properly collocated as Navy Department records.

- **Solution:** Search for Readex Serial Set records with the publication date of 1947 and before and globally change all $a Dept. of the Navy. in the Readex records to $a Navy Dept.

5.9.4 United States Presidential Entries

The presidential headings were initially analyzed for two reasons. First, they are high profile people. Second, each president includes a personal and corporate entry, so data can be compiled on both author types. After searching all presidential entries, four corporate entries and 12 personal entries were incorrect.

- **Solution:** The records for each incorrect heading were globally corrected.

Table 2
Presidential Personal Headings

Incorrect Heading	***Correct Heading***
Polk, James K., $d 1795-1849.	Polk, James K. $q(James Knox),$d 1795-1849.
Grant, Ulysses S., $d 1822-1885.	Grant, Ulysses S. $q(Ulysses Simpson), $d 1822-1885.
Hayes, Rutherford B., $d 1822-1893.	Hayes, Rutherford Birchard, $d 1822-1893.
Garfield, James Abram, $d 1831-1881.	Garfield, James A. $q(James Abram),$d 1831-1881.
Taft, William Howard, $d 1857-1930.	Taft, William H. $q(William Howard),$d 1857-1930.
Harding, Warren Gamaliel,$d1865-1923.	Harding, Warren G.$q(Warren Gamaliel),$d1865-1923.
Roosevelt, Franklin Delano,$d1882-1945.	Roosevelt, Franklin D.$q(Franklin Delano),$d1882-1945.
Eisenhower, Dwight D.,$d1890-1969.	Eisenhower, Dwight D.$q(Dwight David),$d1890-1969.
Kennedy, John Fitzgerald,$d1917-1963.	Kennedy, John F.$q(John Fitzgerald),$d1917-1963.
Johnson, Lyndon Baines,$d1908-1973.	Johnson, Lyndon B.$q(Lyndon Baines),$d1908-1973.
Nixon, Richard Milhous,$d1913-1994.	Nixon, Richard M.$q(Richard Milhous),$d1913-1994.
Ford, Gerald Rudolph,$cJr.,$d1913-2006.	Ford, Gerald R.,$d1913-2006.

Table 3
Presidential Corporate Headings

Incorrect Heading	***Correct Heading***
U.S. President (Johnson, Andrew)	United States. $b President (1865-1869 : Johnson)
U.S. President (Garfield, James)	United States. $b President (1881 : Garfield)
United States. $b President (1885-1889: Cleveland)	United States. $b President (1893-1897: Cleveland)
United States. $b President (1913-1917: Wilson)	United States. $b President (1913-1921: Wilson)

6. Conclusion

Academic libraries need to maintain core primary source collections in this age of overwhelming information output and accessibility. Offering patrons primary sources with excellent digitization, multiple access points, good indexing, and intuitive navigation is key for consistent patron use. Therefore, Readex's digitization of the *American State Papers* and the *Serial Set* holds great significance for academic library faculty, staff, students, and researchers. These collections comprise a historical importance and scope in subject matter that many searchers may find surprising. Prior to the availability of digital records, the extensive collection was known by few researchers, and the only way to navigate the collection was through a cumbersome printed index. At CSUL the nearly 14,000 volumes are available in a thorough database with documents digitized with OCR software and accessible through the library catalog with comprehensive MARC records. Researchers are finding more relevant materials than before because the MARC records are now in the catalog and discovery of *Serial Set* materials is possible without deliberately seeking them.

Even though the for-purchase MARC records for the Serial Set are not perfect, they are a much better alternative than not having them at all or for the library to single-handedly catalog hundreds of thousands of individual items (capturing all of the pertinent keywords) in order to provide a valuable access point. The 400,000 Readex-provided MARC records give article level access to patrons searching the library catalog. Although they contain certain problems, the records are extensive and include great detail. Each record includes ample description in note fields supplemented by Readex Thesaurus subject headings. This metadata can be used by patrons through keyword inquiries, since it is not accessible through exact searches.

The record problems that impact patron accessibility can often be corrected through automated processes. For problems that need manual alteration, there are adequate solutions to finding and making the changes, although these laborious and time consuming solutions will need local definition. With some enhancements (outlined in this article) the MARC records for Readex's Digital Collections provide convenient additional access points to the extensive and essential primary source collections of the *American State Papers* and *United States Congressional Serial Set*.

Acknowledgement

The authors would like to thank August A. Imholtz Jr. for supplying information and journal articles.

References

Association of College & Research Libraries (2006). Readex and Dartmouth partner to digitize congressional holdings. *College & Research Libraries News 67*(1), 9.

Beall, J. (2009) Free books: loading brief MARC records for open-access books in an academic library catalog. *Cataloging & Classification Quarterly, 47*(5), 452–463.

Bernholz, C. D. & Carr, A. G. (2009). The Annual Reports of the Commissioner of Indian Affairs: revisiting the key to the *United States Congressional Serial Set*, 1824-1920. *Government Information Quarterly, 26,* 540-545.

Brown, C. C. (2011). Knowing where they went: six years of online access statistics via the online catalog for federal government information. *College & Research Libraries, 72*(1), 43-61.

Clark, S. M. (1988). Use of the U.S. Serial Set in an academic institution: a collection management tool. *Government Publications Review, 15,* 213-223.

Clausen, B. (2005). Northwestern University Library's experience and decision. *Government Information Quarterly, 22,* 738-742.

DeLong, S. (1996). What is in the United States Serial Set? *Journal of Government Information, 23*(2), 123-135.

Finn, M. (2009). Batch-load authority control cleanup using MarcEdit and LTI. *Technical Services Quarterly 26*(1), 44-50.

Frachtenberg, Leo J. (1920). Alsea texts and myths. House of Representatives. 64th Congress, 2nd session. House Report No. 1023 (*Serial Set* 7438). Washington, DC: Government Printing Office.

Hardenbrook, J. A. (2006). American State Papers, 1789-1838. *Choice Reviews Online, 43.* Retrieved from http://www.cro2.org/default.aspx?page=reviewdisplay&pid=2935540

Harrison, James L. (Eds.) (2010). *100 GPO years, 1861-1961: a history of United States public printing*. Washington, D.C.: U.S. Government Printing Office.

Hartley, J. (2009). Using the Congressional Serial Set for genealogical research. *Prologue, 41*(1). Retrieved from http://www.archives.gov/publications/prologue/2009/spring/congressional-serial-set.html

Imholtz, A. A., Jr. (2008). The American State Papers. *DttP: Documents to the People, 36*(1), 18-21. Retrieved from http://wikis.ala.org/godort/images/5/54/DttP_36n1_web.pdf.

Imholtz, A. A., Jr. (2003). The printing and distribution of the Serial Set: a preliminary contribution to 19th century congressional publishing. *DttP: Documents to the People, 31*(1), 8-17. Retrieved from http://wikis.ala.org/godort/images/a/a6/V31n1_2003Spring.pdf.

Information Today, Inc. (2003, September). Readex will begin digitizing the Serial Set. *Computers in Libraries,* 48-9.

Information Today, Inc. (2004). Readex creates Archive of Americana. *Information Today 21*(11), 31-2.

Kellsey, C. (2002). Cooperative cataloging, vendor records, and European language monographs. *Library Resources & Technical Services, 46*(3), 105-110.

Kellsey, C. (2001). Trends in source of catalog records for European monographs 1996-2000: a preliminary study of Italian monographs. *Library Resources & Technical Services, 45.3,* 123-126.

Larsgaard, M. (1965). Beginner's guide to indexes to the nineteenth century U.S. Serial Set. *Government Publications Review, 2,* 303-311.

Martin, K. E. & Mundle, K. (2010). Notes on operations cataloging e-books and vendor records: a case study at the University of Illinois at Chicago. *Library Resources & Technical Services, 54*(4), 227-237.

Meister, M. & Collins, K. (2005). University of California experience and decision. *Government Information Quarterly, 22,* 742-747.

Miller, A. E. (2005). Digital U.S. Congressional Serial Set collections—reviews of the Readex and LexisNexis products. *DttP: Documents to the People, 33*(2), 37-40. Retrieved from http://wikis.ala.org/godort/images/c/c4/Dttp_v33n2.pdf

Millwood Group Corp. (2006). Readex, Dartmouth partner to digitize publications. *Advanced Technology Libraries, 35*(1), 7.

Mouw, J. R. (2005). Editorial: Library use of vendor-supplied catalog copy. *Library Collections, Acquisitions, & Technical Services, 29,* 235-237.

Mugridge, R.L. & Edmunds, J. (2009). Using batchloading to improve access to electronic and microform collections. *Library Resources & Technical Services, 53*(1), 53-61.

Naylor, C. (2005). Those elusive early Americans: public lands and claims in the *American State Papers*, 1789-1837. *Prologue, 37*(2). Retrieved from http://www.archives.gov/publications/prologue/2005/summer/state-papers.html

New Mexico. (1852). New Mexico: Letter from the Secretary of the Territory of New Mexico, transmitting copies of the acts, resolutions, &c., of that territory. January 9, 1852. House of Representatives. 32nd Congress, 1st session. House Report No. 4 (*Serial Set* 652). Washington, DC.

Norelli, B. (2005). U.S. Congressional Serial Set (1817-1980) with American State Papers (1789-1838)-Readex Digital Edition. *Microform & Imaging Review, 34*(1), 51-53.

Quinn, A. C. (2005a). Cataloguing the Congressional Serial Set. *Cataloging & Classification Quarterly, 41*(1), 183-199.

Quinn, A. C. (2005b). Digital Serial Set [Essay Reviews]. *Government Information Quarterly, 22,* 737-738.

Readex. NewsBank. *U.S. Congressional Serial Set, 1817-1980: Overview.* Retrieved from http://www.readex.com/readex/product.cfm?product=9

Ross, R. A. (1994). Using the *U.S. Congressional Serial Set* for the study of western history. *The Western Historical Quarterly, 25*(2), 208-213.

Rossmann, B. (2004). New options for the Serial Set. *DttP: Documents to the People, 32*(1), 6-7. Retrieved from http://wikis.ala.org/godort/images/2/26/Dttp_v32n1.pdf

Sanchez, E., Fatout, L., Howser, A., & Vance, C. (2006). Cleanup of NetLibrary cataloging records: a methodical front-end process. *Technical Services Quarterly, 23*(4), 51-71.

Santos, S. (2005). Readex *U.S. Congressional Serial Set*, digital editions vs. LexisNexis *U.S. Serial Set* digital collection. *Arkansas Libraries, 62*(1), 26-28.

Stuart, L. M. (2010). The U.S. Congressional Serial Set, 1817-1994. *Choice Reviews Online, 47.* Retrieved from http://www.cro2.org/default.aspx?page=reviewdisplay&pid=3554169.

Thomas, M., Piscitelli, A., & Rholes, J. (1994). Security and preservation of the U.S. Congressional Serial Set. *Journal of Government Information, 2*(4), 351-366.

Treaty with the King of the Belgians. (1865) Treaty with the King of the Belgians. Message from the President of the United States, transmitting a copy of two treaties between the United States and the King of the Belgians in relation to the Scheldt dues, &c. January 9, 1865. House of Representatives. 38th Congress, 2nd session. House Report No. 19 (*Serial Set* 1223). Washington, DC.

Wu, A. & Mitchell, A. M. (2010). Mass management of e-book catalog records: approaches, challenges, and solutions. *Library Resources & Technical Services, 54*(3), 164-174.

Xiong, J. A. (2004). Readex U.S. Congressional Serial Set, digital edition and LexisNexis U.S. Serial Set digital collection. *The Charleston Advisor, 5*(4), 5-12.

Zink, S. D. (1986). CLIO's blindspot: historians' underutilization of United States government publications in historical research. *Government Publications Review, 13,* 67-78.

Vitae

Daniel Draper is an Assistant Professor, Metadata Librarian / Database Management Specialist at Colorado State University.

Naomi Lederer is a Professor, College Liaison Librarian at Colorado State University. She is the author of *Ideas for Librarians Who Teach: With Suggestions for Teachers and Business Presenters* and a number of journal articles.