Managing Serials in a Large Digital Library: Case Study of the UNT Libraries Digital Collections

Abstract

The Portal to Texas History (http://texashistory.unt.edu) and the UNT Digital Library (http://digital.library.unt.edu) contain more than 445,000 items collectively, hosted by the University of North Texas Libraries. These collections include a variety of serial publications such as newsletters, magazines, scholarly journals, annual reports, and yearbooks. Many of these serials are key resources for the user groups of both repositories. In this paper, we discuss the importance of serials within our collections, some of the challenges, the standards we use, and how we leverage metadata to facilitate access to serials for diverse global users.

Keywords: digital libraries, digitization, serial collections, metadata

Introduction

This paper summarizes a presentation of the same title, given at the NASIG 2014 conference in Fort Worth, Texas¹.

The University of North Texas (UNT) Libraries host The Portal to Texas History and the UNT Digital Library digital collections, which together contain more than 445,000 unique digital items. The Portal to Texas History comprises items about Texas history or owned by Texas institutions; the physical collections are owned by over 200 partners across the state, including museums, universities, public libraries, genealogical and historical associations, other organizations, and private collectors. The UNT Digital Library contains items created at or owned by the university, which support the educational interests and needs of the university community.
Although serials represent an important segment of materials, they can be challenging to describe and manage. Digitally-produced serials require less effort to process and archive in a digital library than scanning hard copies, but without an automated mechanism for obtaining copies, or ensuring that the digital collection remains up-to-date as new issues come out, collection maintenance can become complicated. Currently, staff members are responsible for tracking down versions of new issues for ongoing titles in the UNT digital collections, but that may not be a reliable, sustainable, or scalable long-term solution. Additionally, every collection has peculiarities related to cataloging, and serials are no exception. This paper will discuss the importance of serials within our digital collections as well as the standards we use for representing serials and how we leverage metadata to provide better access to serials for our many users.

Serials in UNT Digital Collections

Within the UNT Digital Library, serial publications often represent the work or research interests of UNT faculty members and directly aid their efforts to promote and conduct various kinds of research. For example, we host several publications from local and national organizations, such as the JAC (Journal of Composition Theory), the Journal of Schenkerian Studies (a UNT College of Music publication), the Environmental Ethics journal, and the Southwest Retort (a bulletin produced by the Dallas/Fort Worth Section of the American Chemical Society). Several of these titles belong to under-represented areas of study within their respective fields. Additionally, we host serial publications that promote the university identity, such as the UNT alumni newsletter, UNT yearbooks, annual reports, and departmental newsletters.

Similarly, the Portal contains significant resources representing the heritage of Texas. In particular, we have multiple genealogical newsletters from various organizations across the state, since many researchers use the Portal to locate primary sources and information related
to genealogical work. Additionally, journals from the Texas Historical Commission, the Dallas Historical Society, the Texas State Historical Association, and others help to promote historical research. Other serials simply represent cultural, scholarly, and creative contributions that would not be as accessible otherwise, such as school literary journals, yearbooks, and alumni magazines, as well as ArtLies magazine, Scouting Magazine (the official Boy Scout publication originally for scout leaders), the Journal of the Effective Schools Project, Texas Trends in Art Education (a peer-reviewed journal), and the Cross Timbers Business Report. Without the digital collections in the Portal, the important materials within these serials would be difficult or impossible to access.

Locating Serials

In the UNT digital library system there are several ways to browse, or “Explore” items. One easy way to find many serials is to browse by resource types, which also makes it relatively efficient to gather statistics on many of our serials; as of May 3, 2014, the digital collections included:

- 177,408 newspaper issues and more than 1.6 million pages of Texas newspapers; these are largely excluded from this paper since newspapers are handled by a separate department.
- More than 900 yearbooks from 22 different partner institutions; most of the yearbooks are full runs or near-full collections of yearbooks from various schools.
- 6,033 “Journals, Magazines, and Newsletters” that encompass a range of items, such as scholarly journals, magazines, bulletins, newsletters, and other general periodicals.

However, some serials pose problems when using this method of tabulation since there is no “type” that encompasses all serials. In particular, we have many items classified as “books” or
“reports” that are periodicals, such as annual reports, government documents, technical reports, annual program books, and budgets.

Items are also browsable by the partner institution that owns them (e.g., all serials from a particular university) or by collection. In our system, a “collection” is defined as a group of roughly 50 or more items related by a topic, donor, funding source, or other defining criterion. Often, full runs of serials are also grouped into a collection and contain all of the issues across title changes. Each of the collections also has a description with an “About” page, that functions somewhat like a serial record in our system. The description gives an overview of all items in the collection (or the full serial run) and keeps a running list of the items, with the most recently added items first; this can help users see if any new volumes/issues have been added, or whether the collection is relatively static.

From a search results page, all of these browse options -- type, collection, and institution -- show up in facets for narrowing results. Additionally, there are other facets for options such as date, location, and serial/series title. The title facet allows users to find a collection for an entire serial run and then view all of the title changes or limit to a specific title. For example, a UNT Yearbooks collection contains all 98 yearbooks published between 1906 and 2007, but there are four different serial titles during that time, which appear in the “Serial/Series Title” facet.

In an individual item record, users have links out to the entire collection or to “Other items in this serial” to see all of the issues with the same serial title. We recognize that there are opportunities to add functionality for end users viewing serials; for example it may be helpful for users to have the ability to navigate between one serial issue and another through the addition of a feature that links to the next and previous serial issues in the system for a given title. However, this seemingly straightforward feature introduces a number of challenges for implementation, most notably, how to communicate missing issues -- both unpublished issues and issues missing from the holdings -- to the end user. At this time there are no specific plans for additional features, but it is certainly a possibility for the future.
Metadata for Serials

All items in the UNT digital collections use the same locally-modified Dublin Core metadata schema, containing twenty-one possible fields. For periodicals, our general practice is to create item-level records for serial issues. These records are often less specific than the records for other kinds of items in the digital collections, but allow a level of flexibility for describing problems or unique features at the issue level. For example, we can note pagination that continues across volumes, various inconsistencies, and missing pages. We also add notes about special issues or issue themes by adding special titles, display notes, or information in the content description.

Having this level of granularity also provides a way to represent issues that are “exceptions” within a serial. Although it may not happen often, there are situations when an individual issue has significant differences from the rest of the publication. For example, there is an annual student literary magazine published by McMurry University in Abilene, Texas. However, in 1968, the administration felt that the content did not meet their standards and rejected some of the submissions. The students chose to self-publish a purposely-untitled issue rather than a sanctioned, censored version. The untitled issue is still an important part of the Galleon history, and belongs with the rest of the collection, but it would be difficult to represent this unique issue fully in a serial record.

Treating issues as items also allows for more room to collocate like items. Many of the fields in item records turn into links when displayed to the public, allowing users to find other items that have the same subject, creator or contributor, coverage information, etc. We capitalize on this functionality in several ways. For significant events (e.g., world wars, September 11th, etc.), we sometimes add subject headings to individual yearbooks or journal issues that were published at the time and talk about the events. Another common practice in our system is to use series titles to group sub-sets of items, since series and serial titles both
show up in the browsable facet on search pages. For items such as reports or budgets, we add series titles to link together all of the items that specifically discuss a particular fiscal year; similarly we add titles for federal or state legislative sessions to tie together House Journals, Senate Journals, laws and resolutions, committee reports, or other relevant items that happen within a particular congress or legislature. Finally, for some journals and publications that contain submitted articles, we include each of the authors as contributors, with the titles of the pieces that they wrote. The names also become links so that a user could find all of the issues to which a specific author has contributed articles.

**Serial Dates**

One inherent peculiarity regarding serials is their periodic frequency, which can differ excessively among publications and even over the course of a single title. Some dates have easily-standardized formats (years, months, or issuance days), however, serials in particular tend toward seasonal dates that can be more difficult to represent consistently. In 2009, we switched to the Library of Congress Extended Date-Time Format (EDTF) as our date format of choice. The significant strength of the EDTF is that it provides standardized, machine-readable formats for a range of complex dates that are not part of other date/time standards. For serials it has been particularly helpful, as it allows for a way to represent seasons using a season code in place of the month (e.g., 2014-21 for spring 2014).

Although the EDTF ensures that we represent these dates consistently, it does introduce a new complication, since the average user is not going to recognize or understand a season code, or some of the other complex date strings. However, by using a machine-readable format, we have also been able to introduce date normalization for the user display. While it has not been implemented for all dates, simple dates and seasons are automatically translated into full written text in item records. So, 1981-03 displays as “March 1981” and 2009-24
displays as “Winter 2009.” The EDTF does not fix all date complexities, but it has made dates for almost all serials easy to enter in a consistent format.

**Issues and Parts**

Although we generally handle all items and metadata as consistently as possible, in some specific collections there are exceptions to the rule regarding issues as individual items. For example, the Texas State Historical Association (TSHA) requested that we treat issues of their quarterly journal at the volume level, with one record for each year. Although it groups issues in a way that may be useful in some respects, it also means that the volume-level records are much less robust than issue-level records would have been.

Alternatively, we occasionally duplicate material to represent it more appropriately. One case is the *Hexagon* journal, published by a national chemistry fraternity. A professor at UNT has a series of articles in the journal, documenting his research to track down the locations and history of the original discovery of each element in the Periodic Table. He obtained permission for us to digitize the entire issues of the journal in which his articles appear, but we also wanted to appropriately represent his individual work in our Scholarly Works institutional repository. So, we duplicated the pages containing specific articles with article-level records, which is common for other articles that we receive from our faculty for inclusion in the institutional repository.

Item-level records also become complicated when serials have multiple parts. For example, some newspapers have inserted materials, such as the UNT “Holidaily,” a holiday magazine-type insert added to the UNT student newspaper during winter. In this case, because the *Holidaily* is a regularly-occurring supplement with a separate serial title, we chose to treat it as a separate (but related) serial rather than leaving it attached to the newspaper issue in which it was inserted.

In some other cases, the relationship among pieces is less consistent or well-defined. One of our collections is a near-full run of *Scouting* magazine, a serial published by the Boy Scouts
of America for scout leaders. During a span of years around the 1980s, the magazine included various inserts that appear to have been pull-out pamphlets, generally containing activities for specific age groups. In those cases, multiple pamphlets were often interleaved in the center of the main publication, and there was little consistency in the content and regular inclusion of various titles. For those items, we decided to preserve the original integrity of the item as it was published, since separating pieces would be difficult, potentially inaccurate, and not necessarily add value. However, most of the inserts are represented through series titles and other notes in the item records.

In one highly unusual case, two separate serial titles merged into a single printed publication, but maintained separate titles and issue numbering. We were able to represent both titles in the item records with multiple titles and identifiers.

**Access and Usage**

Although we are committed to open access whenever possible, there are items in the UNT Digital Library that require restrictions due to copyright, licensing, or privacy concerns, while still representing important research materials for the university community. For this reason, we have mechanisms in our system to manage access at multiple levels.

Almost all of our serial materials are openly accessible, however, we have at least one journal managed by a UNT professor that is restricted to the UNT community. This means that anyone can search or view the metadata, but only someone with current university credentials (faculty, staff, or students) can see the item pages.

We also have the option to place items under embargo, when necessary. Most often, this is used for items in our institutional repository that have publisher restrictions. However, several of the journals in our Digital Library have temporary embargoes (usually one to five years) on the most recently published issue(s). In these cases, anyone can search or view the metadata, but no one has access to read the issue until the pre-set date is passed and the embargo is
lifted. This allows the publishers to archive issues as they come out (and make them findable
next to older, public issues), but continue to sell volumes without competition. In most cases,
we try to include contact information in the metadata for users who wish to obtain a copy.

Despite the various complications with access, we do know that people are finding and
using our serials. We track usage statistics at the item level, collection level, and partner
institution level; they are all viewable on the sites. In some cases, journals and other serials are
gaining much more usage simply by being available online. Of course, items such as
newspapers that were only accessible as microfilm have become searchable and easier to
use. But there is also a rise in usage for specialized serials. For example, the *Theoria* journal,
published through the UNT Press, contains articles related to the history of music theory and
usually averages around 100 copies sold per annual volume. Since historic volumes started
going online in November 2013 (roughly eight months), the collection has had over 1,000 uses;
some individual volumes have already registered more than 100 uses, surpassing regular
readership.

Conclusion

Serials represent a specific subset of items in The Portal to Texas History and the UNT
Digital Library, which occasionally have specialized metadata and collection management
needs. In general, we believe that serials are well documented and have useful features in the
UNT digital collections; statistics show that users are finding and reading items from those
collections. However, the UNT Libraries recognize the importance of these materials, and of the
need to continue building serial collections. Additionally, finding ways to learn more about how
end users want or need to interact with serials will allow for future improvements in access and
functionality whenever possible. These measures will help the UNT Libraries to continue
promoting usage of current and new serial collections, making more resources available to
digital users.

Notes

   http://digital.library.unt.edu/ark:/67531/metadc287022/

   http://digital.library.unt.edu/ark:/67531/metadc174739/

Contributor Notes

Hannah Tarver is the Department Head, Digital Projects Unit at the University of North Texas Libraries

Laura Waugh is the Repository Librarian for Scholarly Works at the University of North Texas Libraries

Daniel Gelaw Alemneh is the Digital Curation Coordinator for UNT Digital Libraries at the University of North Texas Libraries

Mark Phillips is the Assistant Dean for Digital Libraries at the University of North Texas Libraries