

# Dean's Innovation Grants

Final Report

**Date:** September 20, 2017

**Project Title:** Using technology to enhance discovery, access, and use.

**Project Team Members:** Kevin Yanowski (PI), Roberta Sittel, Betty Monterroso

## Project Purpose

This project wanted to demonstrate that discoverability and use of items within the UNT Digital Library would increase if those items were included in the UNT Library Catalog. This project also wanted to show that linked data would enhance digital items, specifically government documents.

## Project Activities

1. The team assessed various government document collections to determine their record quality and their broad appeal to a wide variety of users.
  - a. We wanted decent quality digital library records to ease conversion from a modified Dublin Core scheme to an AACR2/RDA hybrid MARC record.
  - b. We also wanted a collection that would have wide appeal across many subjects so that we could get as many usage statistics as possible during our period of data collection.
  - c. We ultimately chose the Congressional Research Service Reports (CRSR) collection because it had good metadata quality, covered a wide variety of subjects, and included documents concerning current events.
    - i. <https://digital.library.unt.edu/explore/collections/CRSR/>
2. The team harvested the CRSR metadata from the digital library using the Marc Edit software. 38,764 records were harvested.
  - a. We then used Marc Edit to select half of the CRSR records at random for our project and left the other half of the records in the digital library only.
    - i. 19,382 records were selected, converted, and edited into an AACR2/RDA hybrid MARC record.
    - ii. 19,382 bibliographic and item records were loaded into UNT's Library Catalog.

3. The team used the UNT Digital Library Google Analytics account and consulted with Jason Thomale and Will Hicks about the best way to track data on the CRSR records.
  - a. The team created a “Campaign ID” in Google Analytics and appended that ID to all CRSR URLs in the UNT Library Catalog.
    - i. This ID allowed the team to track CRSR record use only.
4. Once the records were loaded, the team tracked the use over the course of four months.
  - a. May 1, 2017 – August 31, 2017
5. The team also researched the use of linked data and Bibframe for government documents.

## Budget

The team spent the majority of the budget on part-time student employment funds for our student assistant, Paige Testerman.

## Project Outputs/Outcomes

### Part 1: UNT Digital Library record inclusion in the UNT Library Catalog

1. Using Google Analytics, the team was able to investigate use data and make inferences about user behavior. The team compared user behavior and engagement with the whole of the UNT Digital Library and all of the CRSR Reports, to that of the DIG record set. (Appendix 1) For the purpose of this grant, the team considered time spent on a page, pages per session, and average session duration as key data points.
2. What is significant are the data points “pages per session” and “average session duration,” especially for new visitors. The longer a session duration means that visitors coming to the collection from the library catalog have a more specific research need and thus spend more time engaging with the material. When available, the bounce and exit rates also indicate visitors entering the collection through the catalog are more deliberate in their engagement as they do not exit the resource as quickly and tend to spend more time navigating within the resource.
3. This behavior appears to indicate that creating catalog records of UNT Digital Library material is worthwhile, because it directs motivated, university-affiliated users directly to the content.

### Part 2: Linked data and Bibframe for government documents

1. One potential benefit of using Bibframe/linked data is that it can create a virtual experience similar to browsing library stacks—and that those virtual “stacks” can include not just one library’s holdings, but many libraries’ holdings.
2. In the case of government documents, linked data has the potential to help users see connections between government agencies. Linked data can also facilitate understanding about the implementation of legislation, from the introduction of bills in Congress, to the publication of reports and hearings, to regulatory and administrative

aspects of laws, and finally, to case law. In addition, linked data can lead researchers to related resources that they might not readily find in the physical stacks of a government documents collection, such as visual artwork, music, literature, or dramatic works.

3. One challenge that the team discovered with linked data and government documents is that many items remain uncataloged and undiscoverable. The metadata needed to create the much-anticipated links does not yet exist in any form in many records.
4. The primary result the team found with this project is that the time and cost for converting existing MARC records into a Bibframe/linked data ready format is much less than expected. We are also encouraged by the fact that Bibframe should reduce the overall time and costs associated with authority control work and cataloging in general. This will increase the likelihood that more libraries could to produce the metadata needed and take advantage of linked data in records.

### Other Results (if applicable)

Not applicable

### Anecdotal Information

The team found that when we worked with either cataloging digital library records or converting records to Bibframe, a lot of cross-departmental cooperation occurs. Because of this, the team recommends establishing connections early and starting conversations as soon as possible. This will cut down on wasted time due to one group waiting on another, especially when nearing a strict deadline.

### Best Practices

The team would love for other libraries to look at the possibility of including digital library records in their catalog and to look at the possibility of using linked data and Bibframe.

We present the following best practices based on experience with this project:

- Research and know what groups within an organization might be involved (or want to be involved) before the project begins.
- Have clear, established, and reasonable outcomes and share them with all parties.
  - Example: The team wants to track specific data using Google Analytics, is this possible?
- Have clear reasoning for why the team would like to establish this type of project.
  - Some of the work may go against ideas from other members of the larger organization, but if the team has justification, it can provide specifics to any dissenting party.
- Be ready to present and justify the data and the project, no matter the outcome.

## Program Continuity

The team would like to see the results of this project used as a catalyst to develop a larger and longer-term project that will see UNT Digital Library records included in the UNT Library Catalog.

Appendix 1  
 Google Analytics Data Chart  
 Date Range: May 1, 2017 – August 1, 2017

Collection	# of Sessions	% New Sessions	# of New Users	% Bounce Rate	Pages per Session	Average Session Duration
All DIG Records	39	58.97%	23	17.95%	4.26	0:05:27
All CRSR Records	12,605	85.28%	10,750	50.45%	3.02	0:02:40
All Digital Library	195,733	80.83%	158,211	53.91%	3.34	0:02:40
All DIG Records New Visitor	23	100.00%	23	17.39%	4.48	0:07:07
All CRSR Records New Visitor	10,749	100.01%	10,750	52.59%	2.67	0:02:01
All Digital Library New Visitor	158,137	100.05%	158,211	55.98%	2.95	0:01:58
All DIG Records Returning Visitor	16	n/a	n/a	18.75%	3.94	0:03:04
All CRSR Records Returning Visitor	1,856	n/a	n/a	38.04%	5.04	0:06:23
All Digital Library Returning Visitor	37,596	n/a	n/a	45.23%	5.01	0:05:40