**Funding Research Opportunities Grant (TS/OBS) Report – 2020**

Grantee: Jesse A Lambertson
(Metadata / Digital Resources Librarian – D'Angelo Law Library (UChicago)

**BIBFRAME for Metadata Futures: the Funding Research Opportunities Grant (FROG) Report 2020**

First, I want to thank the grant committee and the persons who agreed to be my references for this grant proposal, Pat Sayre-McCoy & Thomas Dousa. Thank you.

In my work under this grant, I wrangled a few necessary resources I needed that my institution did not own, a survey platform, and a bit of technology to explore the technology behind some of the platforms and their development. I want to thank the committee for that support. This is my general Summer 2020 write-up to the FROG committee based on my responses to the survey I sent out via all the normal library and technical services lists as well as some explanation of the technical operations of the BIBFRAME editor for Linked open Data (LoD) as forked and implemented by LD4P under its grant and the work of the cohort libraries. This was my promise in the original proposal, to generalize comments from libraries about their plans (or lack thereof) to implement BIBFRAME praxis via usage of the LD4P Sinopia Editor and some of the technical requirements for that system to operate and move into production level tasks.

2 Caveats to this report.

1. I already have been adding content to TSLL based on my work in this area, and the next issue of TSLL (September 2020) will feature more.

2. I will eventually be putting together a series of blog posts from a combination of survey results and the technical explorations. Some of these will be added to my own website, but, if TS or OBS wants a blog post from me as well, please ask.

**A. Monies**

Money put to use for the report and future pieces of writing (receipts sent upon request):

1 Purchased SurveyMonkey membership to better work with data received
2. Registered for ALA 2020 (and then moved that registration to 2021 since it was effectively a reduced conference)
3. Purchased a few books that went 'behind the scenes' of JSON.
4. And, finally, the SubtimeText license to look more at code things (has enhanced editing capabilities)

**B. Survey generalities**

The survey itself went out to the majority of, what I thought to be, all the relevant listservs in technical services land, with a target of those institutions such as mine that are active in LD4P processes and Sinopia development.

These included, but were not limited to:

PCC main list
PCC's Linked-Data list
Autocat
BIBFRAME list
Code4Lib
AALL's Technical Services Forum in the community section of their website

A few of the questions were:

Q3: "If you are a metadata or collection services manager, what are your plans for training folks in your department on BIBFRAME?"

Q4: "Please describe your understanding of current cataloging standards: RDA, the LRM, and how you plan to implement these standards in BIBFRAME? What challenges do you foresee in implementing these standards in this new schema?"

Q6: (specifically targeted to those institutions who are working with LD4 stuff): "Please describe your experience in LD4P (if a member of the cohort) and how it relates to your plans for BIBFRAME cataloging in Sinopia?"

There were a range of answers to most of the questions, with some showing general understanding to some showing expert understanding even if their opinions of BIBFRAME as implemented in SINOPIA is a right step or a false one (and there were opinions on both sides of the 'middle' on this question).

I admit out-loud that Q4 was too vaguely worded, and was in fact, too large a question to be asked in a 10-question survey because that one question was really an 'essay' in itself potentially. One of the respondents made that comment to that question.

But generally, I found the range of replies cover three or four positions:

1. Some are vaguely aware of BIBFRAME, know it is coming 'in the future' but as yet had not taken much time of their work days in their departments to train – while some were slightly more aware of this new schema, but knew that the actual change for their library would be tied to whether or not vendors were willing to work and produce BIBFRAME (Resource Description Framework (RDF), Metadata Description Sets (MDSs) instead of 'records' in MARC). I do consider this to be a position from understanding about the deeper workflow changes required for systems and interactions with other platforms.

2. Some were generally aware of the expectations of BIBFRAME as it relates to a schema that would work better for LoD, and that this schema was in heavy development by LC, that Linked Data 4 Production (LD4P) was working on a fork of the LC BIBFRAME editor under the name, Sinopia, and had spent some time reading online documentation and had attended a few webinars or sessions related to BIBFRAME development and its expectations. Most of these respondents also assumed those expectations as realizable and realistic but did not deeply understand the technological elements required to put things in working order.

3. A few respondents knew about BIBFRAME, had heard about LD4P's Sinopia, but as they are not directly involved in its processes, had not much understand of where LD4P was in their own BIBFRAME efforts and their departments had taken little direction toward training in the understandings of BIBFRAME and learning about how it is set up technically (this also overlaps with understanding that vendors must also be on board for large-scale changes to take place across libraries – even if one system, ExLibris' Alma, has some BIBFRAME conversions taking place internally and there seems to be no discussion anywhere about how Alma's internal conversion will match what wither LC or LD4P is doing).

4. A few respondents have much expertise in LC's development of BIBFRAME, as well as LD4Ps efforts in Sinopia and LoD, but think the efforts are mis-directed and will not be realized very well (this was the case for reasons of technical and other reasons). [I was personally very happy to see these very few critical responses and that there are a diversity of opinions 'out there.'] On the flipside, there were a few respondents who take a positive approach with similar levels of expertise.

*And an unofficial* 5. I also received a few responses from folks that were e-mailed with specific questions because of their place in high-level development for LD4P and its engagement in BIBFRAME as LoD schema of choice in their Sinopia Editor. The folks who responded to this specific set of questions: Michelle Futornick (LD4P Project Manager who works at Stanford, Justin Littman (developer at Stanford), Jeremy Nelson (developer at Stanford). I had hoped to hear from a couple folks at Library of Congress, folks working on their BIBFRAME editor – but they did not respond.

**C. Technical Configurations**

The really in-depth analysis of BIBFRAME and Sinopia will come out later in going articles to TSLL, these blog posts I mentioned above, and perhaps an article or two submitted for publication later in the year (or early next year), but I do want to explore some of the technical differences in metadata as produced in RDF vs metadata produced as a 'record' we now see in MARC.

I want to shed some light on the processes underway, how I have been using BIBFRAME as implemented in Sinopia, and some of the technical things that make this SO different from MARC cataloging.

First of all, much of the work in LC's BIBFRAME editor and LD4P's fork of that editor in Sinopia is open for public viewing on their respective GitHub profiles
LC's BIBFRAME editor: https://github.com/lcnetdev/bfe
LD4P's Sinopia editor: https://github.com/LD4P/sinopia_editor

These editors both work in the browser and thus have no client like one would have in the form of say Connexion of Marc Edit. Both of these editors are general RDF editors and can, theoretically adapt to or edit RDF data coming from different sources. When we say, RDF, we mean RDF triples in the form of subject, predicate, and object, i.e., "Creator Agent" "Has label" object, "Jesse Lambertson" for this report (as an example) – in the form, ideally, of triples of URIs in which each item of the triple is an URI (which is what we would expect from a LoD environment). There can be 'literals' in the 'object,' but even this, when minted in the Editor, becomes a triple within the editor and is actually given a URI.

These triples are then ‘serialized’ as JSON when exported or communicated to a different system. What this really means is that the rules for meaning, manipulation, display, etc., are NOT written within the serialization, but are tied to the object status (i.e., once they are de-serialized). The rules for those things are configured external to the serialized data when it is being exported or sent to some other system for some purpose or transformation. This fact makes JSON, a current standard for serialization, very flexible for adoption by lots of systems in lots of contexts. JSON is kind of, and I do mean, ‘kind of’ what XML used to be, but is even simpler because it lacks all the tags we find in XML. Now, some systems can work with RDF in XML, but the Sinopia editor is working in JSON and exports all the RDF as JSON-LD (JSON – Linked-Data). The interaction with the metadata is just different than it is in our current MARC-based metadata. I don’t want to get too much in the weeds in this report, but I do want to sort of poke behind the curtain a wee bit to see what is happening there and how BIBFRAME, or its fork by LD4P in Sinopia, is changing the nature of structured metadata production that is LoD in nature and structured in ways more compatible with current web technologies.

I hope to write something further in the future regarding:

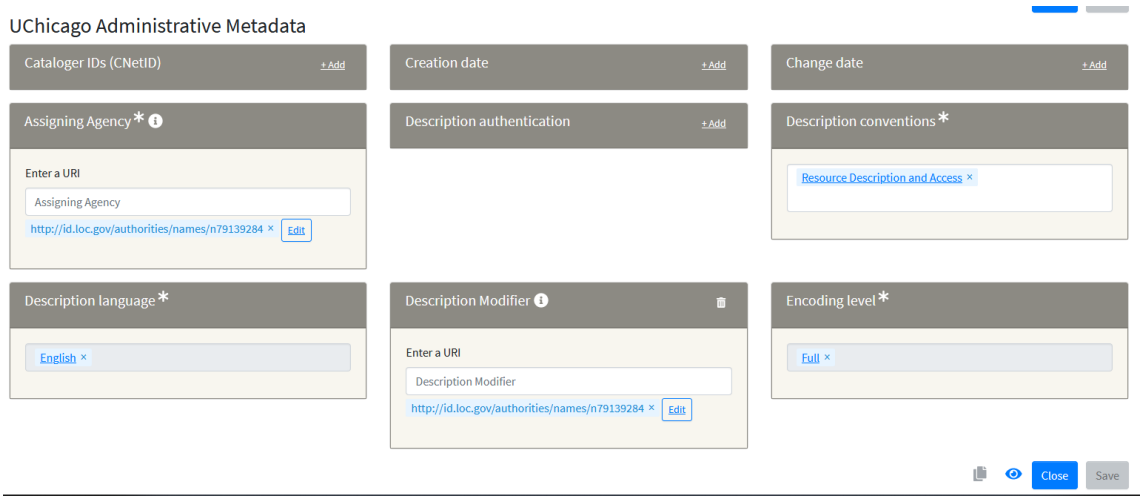
1. BF2MARC Conversion
2. MARC2BF Conversion
3. Conversion of data produced in Sinopia
4. RDA in BIBFRAME
5. Issues of LRM in BIBFRAME
5. Other topics as time allows (say, for instance, better compatibility with current web technology)

For an earlier Description & Entry column in TSLL, I wrote about something fascinating that involves the work-in-process of Administrative Metadata as connected with cataloguing in Sinopia.

This is the metadata that defines ‘who`’ did ‘what’ ‘when.’

We have become accustomed to personalized logins that control who can do what kind of change to what kind of data in a system. Some of these logins can tether changes to ‘who’ did ‘what’ and ‘when’ it is done. When I said this is a work in progress I meant it. In the current Sinopia system, very little besides login permissions are associated with a user logged in.

Nothing specifically is set to configure or auto-populate data points about ‘who’ did ‘what.’ On the contrary, a specific template had to be configured to produce some kind of history connecting user interactions with metadata for a resource (with key data being added manually on a per-edit basis)




Theoretically, the Cataloger ID would be auto-populated by the login information, but that is not the case at this point. What we can produce as a default, is that one's local institution is working on the recource in-hand, the source of said institution's URI (in this case, from LC NAF), but also default encoding levels and language of cataloging (in this case English, as is the standard for metadata production in the United States.

Let me quickly show what this configuration of Administrative Metadata looks like from a portable exported form in JSON:


```
{
  "propertyTemplates": [
    {
      "mandatory": "false",
      "repeatable": "true",
      "type": "literal",
      "propertyURI": "http://id.loc.gov/ontologies/bflc/catalogerId",
      "propertyLabel": "Cataloger IDs (CNetID)",
      "resourceTemplates": [],
      "valueConstraint": {
        "valueTemplateRefs": [],
        "useValuesFrom": [],
        "defaults": []
      }
    },
    {
      "mandatory": "false",
      "repeatable": "false",
      "type": "literal",
      "valueConstraint": {
        "valueDataType": {
          "dataTypeURI": "http://www.w3.org/2001/XMLSchema#date"
        },
        "useValuesFrom": [],
        "valueTemplateRefs": [],
        "defaults": []
      },
      "propertyLabel": "Creation date",
      "propertyURI": "http://id.loc.gov/ontologies/bibframe/creationDate",
      "resourceTemplates": []
    },
    {
      "mandatory": "false",
      "repeatable": "false",
      "type": "literal",
      "valueConstraint": {
        "valueDataType": {
          "dataTypeURI": "http://www.w3.org/2001/XMLSchema#dateTime"
```

This is the same data as that above was a screenshot from the browser as interacted with by the cataloguer. But not one important thing: there is NO data in this file defining aethetics, spacing of elements, schemes, etc. ALL those elements are described in the design files. But the serialized JSON here, looked upon from an export from Sinopia, presents the bare structure of the Administrative Metadata and THAT IS IT! I highlight it to signify that our metadata work and the structure of said work, in this new environment is being organized by completely different interactions now.

One thing that must be brought up when we think about how cataloging is determined by technology: there are all new and challenging fronts to this new world and we are aware that integrations with systems and vendor processes MUST be developed. In fact, in my survey, as mentioned above, quite a few people noted the fact that the impetus to adopt and learn BIBFRAME, or jumping into metadata production in Sinopia, is tethered to adoption by vendors, eCIP workflows, and the like. There are A LOT of technical things that have to change for linked-data metadata to be produced and used across systems. Also, just as a note, LD4P (grant 3) is being set aside to ‘close the loop’ and develop bridges and interfaces for Sinopia to export and import data as needed from external sources such as discovery, record conversion processes, and maybe even from vendors (in the form of stub MDSs to begin cataloging).

So, that’s my update generally. I will go further into the weeds later, am expecting to submit a final report n September since the original completion date was the end of September 2020. I look forward to making more progress and learning more as the work carries on.

Jesse A Lambertson
lambertson@uchicago.edu