PCC Working Group on Metadata Application Profiles Report--November 2021

Introduction

The MAPs Working Group was charged in June 2021 to oversee the creation and maintenance of PCC Metadata Application Profiles and cultivate expertise in Metadata Application Profile design and functionality within PCC; collaborate with the relevant PCC standing committees as needed; communicate with and perform outreach to stakeholders (library services platform and discovery service vendors, data remediation services, other creators of profiles in the cultural heritage sector).[1] Initial membership was: TJ Kao (Co-Chair), Nancy Lorimer (Co-Chair), Everett Allgood, Manon Théroux, Paloma Graciani Picardo, Haiqing Lin, Sophie Dong, Margarita Perez Martinez, Jeannette Ho & Tina Shrader. Everett has since had to drop out due to family illness; Charlene Chou has kindly offered to step in in his place for now.

Our first task, as we understood it, was to prepare the textual monograph and serial templates in Sinopia for use by PCC libraries as base metadata application profiles for both testing the templates and for broadening experience in cataloging using linked data, particularly BIBFRAME-based descriptions. The Working Group has thus far only partially fulfilled this goal, and thus is somewhat late in completing our first deliverable. TJ and I are confident, however, that we will catch up in the near future.

General Report

The membership of the working group is purposely made up of a broad range of metadata expertise. To make good use of that range, we have tried to approach the MAPs from a few different perspectives, while starting to broaden the general knowledge of Sinopia templates and their MAP roles. A template is a good basis to understanding MAPs, in that they illustrate in a practical manner, the restrictions and choices that a MAP will include. To start, Nancy gave an introduction to template creation in Sinopia, and made sure everyone opened user accounts. Members also delved into the training modules and videos provided by PCC and the PCC Sinopia Cataloging Affinity Group. We then began to compare the monograph template MAP with that of the BSR, noted errors, inconsistencies in template vocabulary and labeling, and ambiguities in how data should be entered in the template. We continue to collect these observations both on spreadsheets and as questions for easy accessibility to all members as they familiarize themselves with use of the templates. We also have begun to explore ways to integrate the BSR into the templates themselves.

[1] For the full charge see https://www.loc.gov/aba/pcc/taskgroup/metadata-application-profiles-wg.pdf

While we have collected a rich collection of issues and changes, we have not as yet adjusted the templates to reflect the new additions. There are two primary reasons for this: 1. Close to the start time of our working group, the Library of Congress published an update to BIBFRAME, removing domains and ranges for several properties, creating inverse properties, and introducing some new properties, particularly for serials. We have not had the time or opportunity to study these changes more and determine what effect they might have on the PCC templates; 2. Sinopia developers began a double work cycle at the end of August. This means that they are working on enhancements to Sinopia until mid-December, with only a one week break in mid-October. Some of these changes are major--the introduction of permissions, addition of datatypes and validation, the ability to add pairs of URIs & labels that are not in the lookup, RDF to MARC conversion and export to a catalog, and various other enhancements, several of which were recommended by PCC. Because many of these enhancements have a direct impact on the functioning of the templates, and the templates themselves are being used in testing, it has not been the best time to start adding a lot of things. Some interim changes have also made template creation somewhat confusing, a situation that is not so good for beginners and even the more experienced members. Some of our changes are also waiting on enhancements that have not yet been completed (multiple remarks; datatypes).

Despite all this, we would like to emphasize that the monograph templates do work, and have mostly continued to work during this work cycle. PCC template-based monographic descriptions can also now be exported as MARC for use as operational records in an ILS (currently Symphony only, though working on FOLIO). In that way, we have fulfilled our initial requirements for monographs, though not yet to our satisfaction. The serial template is much less tested and needs more attention.

Outreach

Outreach is another aspect of our working group. While we do not yet have a direct effort in this, we have interacted a fair amount with members of the PCC Sinopia Cataloging Affinity Group and have begun to explore ways for feedback from the group to be funneled back to the Working Group. Also, through connections with other projects, we have asked the MLA Linked Data Working Group, currently led by Kevin Kishimoto, to create templates, based on the PCC monograph template, that will provide the basis for PCC MAPs for music audio recordings and scores. That MLA provides its own MAPs is consistent with the development of the BSR, for which MLA provided the music requirements.

Moving Forward

As some aspects of the Sinopia work cycle wind down, the group will begin enhancing the templates, beginning with the simplest and least intrusive (adding remarks, basic guidance) and moving forward from there. We aim for the template to continue to be fully functional and be able to test an enhanced version before the Sinopia work cycle fully ends. We will also simultaneously create human-readable MAPs, experimenting with a few different options, including a simple text version, as with the BSR, a manual spreadsheet such as that created by

the PCC Task Group on MAPs, and a machine-created readable MAP through the DC Tabular Application Profiles (DC-TAP) developed through DCMI. We will then move on to the serials template.

We will also continue to explore our relationship with the Sinopia Cataloging Affinity Group, finding ways for both groups to thrive and to share our experiences, and to work with the MLA working group, even if this goes somewhat beyond our current charge.

Questions

A number of questions have been raised by us, or asked of us, as we do our work that we now bring to SCS and PoCo.

1. LCSH vs FAST
   When it comes to subject, it is suggested that catalogers use OCLC FAST instead of LCSH post-coordinated form. Sinopia does not handle post-coordination and is unlikely to do so any time soon. It is our understanding that LC is considering the creation of URIs for post-coordinated strings, but it does not seem likely that will help in the short term, since so many would be required to make a difference. Also, do we care about the order of subject headings? Sinopia can instill order, but at the risk of incompatibility down the line.

2. Should we be using the Original RDA or the Official RDA?
   This has been asked of the group and by the group. There are some that think we should be moving now to Official RDA so that our templates, and thus our data, is compatible with the newer version. This is particularly an issue with some specifics of Official RDA, the aggregates model, for instance, in which the modeling between FRBR and LRM are in conflict.

3. As is always the case with PCC, there is the question of how far we conform with what LC does and how much we push for something else if we think it better. Do we try to keep our templates as close to those of LC's as possible? Or do we risk some initial incompatibility to test out alternative models (e.g., aggregates) in our templates? In addition, LC has indicated that major changes need to be tested before they are adopted into BIBFRAME. How can PCC do that and still maintain a standard?

4. We have been receiving requests, particularly via the PCC Sinopia Cataloging Affinity Group for templates for other types of resources. We have already partnered with MLA for some music templates. Should we be expanding our template scope?

5. The templates are currently in the Sinopia stage server rather than in production. This means there is no export of descriptions or any guarantee that the data will be saved over time (in fact, a purge of data was just completed at the beginning of this work

cycle). OCLC and SVDE will need a reasonably sized set of descriptions based on our templates to test loading/processing/reconciling the data within their own environments, an important step in building out the linked data infrastructure. What criteria should we take into account in the decision to move PCC templates to the production server? Do we want to create a PCC dataset in production that will serve as test data for OCLC & SVDE or would we rather an individual institution (using the PCC templates) provide the dataset instead?