PREMIS Editorial Committee Conference Call Notes

6 September 2007

In attendance: Rebecca Guenther, Brian Lavoie, Steve Bordwell, Gerard Clifton (notes).
Apologies: Olaf Brandt, Priscilla Caplan, Bill Leonard, Yaniv Levi, Rory McLeod, Zhiwu Xie.

1. Review of Progress

In general, we are not following the procedure we set out earlier in terms of finalizing changes. We are making good progress (although a little behind where we expected to be), but we have been less formal in following through on some of the steps, especially writing up in the appropriate DD section and format, and updating the 'list of changes' to indicate that items have been finalized.

Finalizing some items may depend on resolving three main issues that have arisen during the review process: 1) extensibility, 2) controlled vocabularies, 3) application of objectCharacteristics (or other units) at Representation level.

ACTION: EC members should each:

- Check the items for which they are responsible from the 'sign-up' list of changes.
- If draft proposals are available, update 'TBA' to link to the proposal.
- If items have been revised, but not yet finalized, indicate progress on the 'sign-up' list.
- If items have been discussed and a decision reached, **write up the change formally in the DD format**. Post the DD section on the EC wiki and provide 'Finalized' links from the list of changes – we need to readily see what has been resolved and what has not.
- [Should resolved items also move over to the PIG site? Or the main LC site?]
- If items are dependent on any of the three main issues, indicate this is in the list of changes.

Aim to make finalized items in DD format able to be readily collated – put them in one place, or consistently linked. An editor (possibly Cynthia Hodgson, NISO) will be engaged to put the final document together.

There will also be additional revisions required, especially where there are new topics introduced or revisions which need to be reflected in several places and not just in the relevant semantic unit.

ACTION: ALL: For each change, consider what other changes are required in: 1) introduction; 2) Data Dictionary units; 3) special topics; 4) examples (revised or new examples to illustrate use of revised/new units). Consider if any new sections are needed – e.g. a new special topic.

We should also begin indicating these changes and constructing new examples.

ACTION: Brian Lavoie volunteered to start looking at the introduction and to make list of sections that need to be changed (e.g. data model). It was suggested that Olaf Brandt might consider providing examples for extensibility.

2. Controlled vocabularies

Brian provided a revised proposal for data constraint container – stripped down to a 'vocabularyIdentifier container' (see http://pec.lib.uchicago.edu:8888/pec/uploads/1/vocabularyIdentifier.doc).
The rationale for the container is to promote interoperability and convergence towards best practice. The container contains:

- a vocabularyIdentifierType (e.g. local; URL);
- a vocabularyIdentifierValue (e.g. http://www.w3.org/TR/NOTE-datetime); and
- a vocabularyIdentifierName (e.g. ISO 8601: Date & Time Formats)

Discussion:

- Usage notes for the vocabularyIdentifier container: May be preferable to say that a repository may or may not record this explicitly, replacing 'It is likely…'
- The 'formatRegistry' example in the rationale may imply an expectation that the vocabulary being indicated by the container is a formal, authoritative source. However, repositories may be using a local vocabulary – it may not be formal, and repositories may not wish to make it a formal, public, maintained vocabulary.
- Repeatability: Is there a need to allow multiple vocabularies to populate the associated semantic unit? Perhaps make the vocaularyIdentifier container non-repeatable.
- Obligation: vocabularyIdentifier may be mandatory in the sense that it should be known, but a repository may not want to be explicit (or to formalize the vocabulary). Are we providing this as a convenience, or is it core (it wasn't considered so first time around). It may need to be made optional, similar to the National Library of New Zealand approach (to Environment), in that we don't know how difficult it may be to implement as mandatory or 'mandatory if applicable'.
- There might also need to be another subunit describing which semantic unit the vocabulary applies to.
- Implementation: Will be difficult to place/indicate in the Data Dictionary – do we need to include it each time for each applicable semantic unit? We might be adding containers for everything. Alternatively, implementation could be described in a special topic. May be relatively simple to implement in XML schema (e.g. as attributes).
- Issues:

- o Infinite regression for vocabularyIdentifierType (which 'should be taken from a controlled vocabulary' – where would this end?). Suggestions: use the most specific term possible, or use an internal/local vocabulary. This could be one vocabulary for which we actually provide a value list.
  However, how useful is vocabularyIdentifierType? It is likely to be either a local value or a URL/URI. It may not add much information to the other units.
  - o As indicated above, how will this be shown in DD context?
- ACTION: Brian Lavoie to do a few examples.
- ACTION: Rebecca Guenther to look at DD and see how this will affect containers (where recommended), and propose some options on how it could be implemented.

3. Dates

Rebecca provided an update on consideration of date formats. The main issue is in not being able to express open-ended dates. Also, there are commonly two alternatives for representing dates – with or without hyphens. The dateTime conventions were only ever written as a W3C note and were not formally standardized – they just began to be implemented.

Rebecca has contacted John Kunze (CDL) about the TEMPER proposal, to see if this can be advanced – e.g. put forward as an Internet Draft – for W3C or ISO or IETF-RFT acceptance.
It may not need to be a formal standard, but it would provide a concrete convention for use in the DD. We can then point to it using the vocabularyIdentifier method – when resolved.

4. Other business

There is a proposal for a BOF session at the DLF Forum in November. It was also suggested that there be a session on the revisions, but it was decided that this may be too premature – we should wait until we are done!
ACTION: Rebecca to circulate the proposal for the BOF session.

Next call: 20 September [revised to 27 September]