PREMIS Editorial Committee Conference Call Notes

19 April 2007

<u>In attendance:</u> Rebecca Guenther, Rory McLeod, Bill Leonard, Steve Bordwell, Olaf Brandt, Gerard Clifton (notes).

<u>Apologies:</u> Zhiwu Xie, Yaniv Levi, Priscilla Caplan, Brian Lavoie.

<u>1. JISC funding calls</u>

Rory reported that the JISC call for submissions for funding on 'significant properties' studies has just closed. However, in general, the EC is eligible to apply for JISC funding, provided that proposals have a UK lead – such as Rory at BL, Steve Bordwell at National Archives of Scotland, or both. The EC may wish to consider bidding in future funding calls. There may still also be a possibility to submit for the current funding call on 'significant properties'.

ACTION: Rory McLeod to:

- find out the response level and status of the 'significant properties' bid, in case EC can still submit.
- discuss with Steve Bordwell whether there might be staff resources available (e.g. via HATII or DCC) for this or future studies.
- keep a standing watch on JISC funding calls and communicate them to EC.

<u>2. Group discussion of "relationship" DD changes</u>

<u>Change 13 – linkingIntellectualEntityIdentifier</u>

Issue: "What is the appropriate interpretation of linkingIntellectualEntityIdentifier: seems like it should be for an abstraction, but DD seems to indicate you can link to descriptive metadata for a specific representation. Is this right?" (http://listserv.loc.gov/cgi-bin/wa?A2=ind0606&L=pig&T=0&P=1990)

Discussion: The link to an intellectual entity could be at various levels: e.g. a link to a catalogue record which could be a surrogate for one representation or many. It need not be restricted to catalogue records alone; any metadata could represent the intellectual entity. The link could also be to a specific representation at any level.

Decision:

Adjust usage notes to indicate that PREMIS is not prescriptive on what linkingIntellectualEntityIdentifier should link to; it is up to the institution to decide the appropriate form representing the Intellectual Entity. The representation can be at different levels.

Change 14 – relatedObjectSequence

Issue: "relatedObjectSequence is mandatory but does not apply to derivative relationships (not mandatory in schema but is in PDD) (APSR)"
(http://www.apsr.edu.au/publications/presta/#appendix7)

Discussion: relatedObjectSequence appears suited to use in describing structural relationships, but does not apply to derivative relationships (although, it was suggested that it may apply for a sequence of objects representing a series of migrations of an object). If it were mandatory, a zero value could be used to indicate 'no sequence', but this still would not seem appropriate to derivative relationships. (This may also be the converse case to Change 21, below.)

Decision:
- Make relatedObjectSequence optional in the schemas (if it isn't already).
- Make relatedObjectSeqence optional in the Data Dictionary
- Include a usage note stating that it should be considered Mandatory for structural relationships, but does not apply to derivative relationships.

Changes 20 & 21 - relatedEventIdentifier

Issue: 20 - "relatedEventIdentification is mandatory but only applies to derivation, not structural, relationships (APSR)." 21- "relatedEventIdentification should not be mandatory."
(http://listserv.loc.gov/cgi-bin/wa?A2=ind0507&L=pig&T=0&P=52)

Proposal: (Bill Leonard) http://listserv.loc.gov/cgi-bin/wa?A2=ind0704&L=premis-ec&T=0&P=1769

Discussion: This issue is similar to that for relatedObjectSequence, but in this case relatedEventIdentification applies to derivative, not structural, relationships.

Decision:
- Make relatedEventIdentifier optional in the Data Dictionary
- Include a usage note stating that it should be considered Mandatory for derivative relationships, but does not apply to structural relationships.

Change 15 - relatedEventSequence

Issue: "relatedEventSequence optional in DD, mandatory in schema." (http://listserv.loc.gov/cgi-bin/wa?A2=ind0602&L=pig&T=0&P=980)

Discussion: This follows on from Changes 20 & 21 and is similar to Change 14 (relatedObjectSequence).

Decision:

- Make relatedEventSequence optional in the schemas. (It is already optional in the DD.)
- Include a usage note stating that it should be considered Mandatory for derivative relationships, but does not apply to structural relationships.

3. Discussion of Change 29 – extension of format information

Change 29

Issue: "Suggestion that all format-related information should be in one place." (http://listserv.loc.gov/cgi-bin/wa?A2=ind0606&L=pig&T=0&P=395)

Proposal: (Steve Bordwell) http://listserv.loc.gov/cgi-bin/wa?A2=ind0704&L=premis-ec&T=0&P=1051

Discussion: The issue is that an implementer wants to record detailed information about a text format which is not currently defined in a format registry for reference. The proposed solutions are to keep the detailed format information locally, referencing it with an appropriate value from a local controlled vocabulary, or to include the format information in a new semantic unit 'formatOtherInformation' within the 'formatDesignation' container unit.

The EC agreed that the kind of format information in question would constitute format-specific metadata, which is outside the scope for PREMIS. Within a METS implementation, the textMD extension schema could be used to record this metadata within a separate techMD section. However, other container formats (e.g. MPEG21-DIDL) might have different means for handling this, so the solution would look different. To be implementation-neutral, the answer would be to use whichever method an implementation had chosen for handling other kinds of format-specific metadata.

The use of a controlled vocabulary which refers to a locally defined format definition or specification could also be a solution, as PREMIS does not specify which controlled vocabulary should be used for 'formatDesignation'; it could be a local one.

The proposed inclusion of an 'other information' unit for format information would probably need some structure below that to be useful (otherwise it may be just textual description), and this would need some form of control.

Decision:

- Specific information about text formats should be recorded in the same way as an implementation currently records format-specific information, which is outside scope for PREMIS. In the present case, the use of textMD within METS was suggested as an example.
- Alternatively, local format definitions recorded outside PREMIS could be referenced using a local controlled vocabulary, as the controlled vocabulary is not specified within 'formatDesignation' units.

ACTION: Steve Bordwell to respond to the original PIG message which raised this query.

4. Discussion of Change 26 – clarification on 'parent/child ' relationships

Change 26

Issue: The use of 'parent', 'child' and 'sibling' for expressing hierarchical relationships in relationshipSubType is confusing, as the distinction between structural (whole/part) relationships and derivative relationships (B derived from A) is not clear. Use of more specific terms defining the relationship is suggested – e.g. has_part, is_part_of, source_of, derived_from. (http://listserv.loc.gov/cgi-bin/wa?A2=ind0602&L=pig&T=0&P=4317)

Proposal: (Rebecca Guenther) http://listserv.loc.gov/cgi-bin/wa?A2=ind0704&L=premis-ec&T=0&P=1287

Discussion: There was discussion about use of 'sibling', with the suggestion that perhaps this should only be used between representations, not objects, (using the example of two objects derived from a common parent, resulting in different 'sibling' representations [in a derivative relationship?]). However, a sequence of images representing pages in the same object would also be regarded as 'siblings' in a structural relationship. Similar issues arise for use of 'parent/child' in distinguishing structural (whole/part) or derivative relationships.
The issue seems to be one of specificity in controlled vocabularies. It should be rolled into wider discussion on that topic.

Decision:

- This issue should be included in wider discussion of controlled vocabularies.

- EC should establish some general principles for controlled vocabularies, then put up some draft controlled lists for comment.
- Comments already exist for this controlled vocabulary (for relationshipSubType) - EC should consider changing the suggested values.

6. Other business

PREMIS tutorial in Albuquerque – Rebecca to present this tutorial with Zhiwu Xie today. Zhiwu's slides on use of PREMIS in MPEG21-DIDL are available on the PIG wiki.

Use of Google Docs for DD revision – In a message to EC, Zhiwu suggested using Google Docs for collaborative editing of revision drafts of the Data Dictionary. From members' previous experience, it was noted that there can be issues for retaining detailed formatting in this environment, so this approach was not recommended for editing the DD.

NEXT CALL: 3 May 2007
9:00am Eastern US, Ottawa
7:00am Mountain time US
14:00 UK
15:00 Europe/Germany
16:00 Israel
23:00 Canberra AU

Rebecca will be on vacation. Brian Lavoie suggested to chair 3 May call.

Summary of Actions from this call:

- **ACTION:** Rory McLeod to:
  - find out the response level and status of the 'significant properties' bid, in case EC can still submit.
  - discuss with Steve Bordwell whether there might be staff resources available (e.g. via HATII or DCC) for this or future studies.
  - keep a standing watch on JISC funding calls and communicate them to EC.
- **ACTION:** Steve Bordwell to respond to the original PIG message which raised the query on detailed text format metadata.