

PREMIS Editorial Committee Conference Call Notes

24 January 2008

In attendance: Rebecca Guenther, Priscilla Caplan, Angela Dappert, Zhiwu Xie, Gerard Clifton (notes).

Apologies: Steve Bordwell, Olaf Brandt, Brian Lavoie, Bill Leonard, Markus Enders

1. Extensibility

Proposal: Rebecca put forward a new proposal for extensibility (see <http://pec.lib.uchicago.edu:8888/pec/43>), drawing on previous discussions (see summary at: <http://pec.lib.uchicago.edu:8888/pec/42>), but simplifying the implementation. The proposal:

- nominates a limited set of semantic units for extension – those where greater granularity is likely to be needed or already available via other schemas:
 - creatingApplication; keyInformation; rights; eventOutcomeDetail
- adopts the principle that extensible elements must be containers
- introduces an <extension> schema element under the containers to handle extension schemas. The element is in schema only and is not included in the DD.
- suggests inclusion of the <extension> element also directly under each Entity.
- introduces an <additionalTechnicalCharacteristics> container semantic unit to accommodate additional format specific technical metadata.

Discussion:

- There was general agreement that the new proposal appears to work well, covering the main needs for extensibility while remaining relatively simple.
- The semantic units already identified are a good start for this revision. With the introduction of the <extension> element, further semantic units could be extended later if the need becomes apparent.
- The semantic units discussed for extension were:
 - CreatingApplication – YES – already noted in the DD.
 - Rights – YES – as previously discussed.
 - keyInformation – YES – just the keyInformation container, not all of signatureInformation, may be replaced by XML Signatures (but check what is covered by this XML schema).
 - EventOutcomeDetail – YES – use Rebecca's solution as is.
- Other units discussed:

- Environment:hardware & environment:software – YES – registries may provide environment information in schema
 - Format – NO – additional technical metadata covered by a separate container, additionalTechnicalCharacteristics (see below)
 - significantProperties – NO – we have included additional flexible structure within the revised container already. This could be extended at a later date when more is known.
- Of the semantic units agreed, only EventOutcomeDetail is an element; all the others are already containers. A new container EventOutcomeDetailInformation is proposed to contain EventOutcomeDetail and the extension element.
- The new container, additionalTechnicalCharacteristics, should stand as a new semantic unit in the DD. It is intended to contain additional, mainly format specific, technical metadata (e.g. JHOVE output). There was a suggestion that it could be just “additionalCharacteristics” and could include other information (e.g. legal notes, preservation policy); however, this kind of information is more the ‘business rule’ type of metadata and is therefore out of scope for PREMIS. It is preferable to specifically limit “additionalTechnicalCharacteristics” to technical metadata (e.g. output from characterization tools – schemas for JHOVE, MIX, NLNZ-extractor etc. – or other sources). It also seems preferable to NOT include additionalTechnicalCharacteristics within the format container, as additionalTechnicalCharacteristics could contain other technical metadata not directly related to format. [Notetaker’s note: I’m not sure we decided where additionalTechnicalCharacteristics should go – just under Object?]
- The <extension> container is used to hold additional or more granular metadata relating to the extensible semantic units. In a relational database, additional metadata can be inserted as required, but in exporting to PREMIS schemas the new container is required to accept the additional metadata. In exchange with other repositories, metadata in extension containers may need to be excluded.
- The <extension> element would appear in the schemas only; it need not appear in the DD (although there was a suggestion that it could be described in a table, like other semantic units, if it makes sense to do so). It should be clearly described in the DD where and how the extension unit should be used.
- Extension units themselves must not be repeated. The main container in which they are used should be repeated instead. (This is an XML requirement.)
- For optional containers, it is allowable to use just the extension container without any other PREMIS units, where they exist (the extension could be used to replace the whole structure of the unit). (If any contained PREMIS elements were mandatory, some

way to declare that the extension contents replaced the mandatory units would be required, but this is unnecessary so far.)

- Where a 1:1 mapping of extension contents to PREMIS units is possible, it could be a best practice recommendation to convert the extension contents to the PREMIS units. (This should not be a forced conversion, however, as this would seem to contradict the notion that these semantic units are optional).
- A general extension unit should NOT be included directly below any of the PREMIS Entity containers, except for Rights, which may be entirely replaced by a different schema.
- Could we allow reference pointers to metadata stored somewhere else (e.g. for technical metadata)? (Whether we require elements similar to METS <mdWrap> or <mdRef> is a schema question. We could define in the schema a content type of 'any' with lax processing to allow either a reference or value – again, a schema question.)

Decisions:

- We adopt the principle that only containers may be extended.
- Extensibility should be applicable to: creatingApplication, rights, keyInformation, eventOutcomeDetail, environment:hardware, environment:software
- An <extension> container should be included in the schema definitions of these elements. It does not need to be included as a unit in the DD, but its use should be clearly described.
- (For optional semantic units) extension containers can be used to replace or supplement semantic units.
- For optional semantic units, the enclosed extension container can be used without other PREMIS units.
- For optional semantic units, where a 1:1 mapping with the extension contents exists, use of the corresponding PREMIS elements is recommended as best practice, but is not required.
- When using the extension container in the schema, implementers must declare the extension schema being used (e.g. via a namespace declaration, URI link, etc.)
- Extension containers may not be repeated. Repeat the enclosing PREMIS container instead (if allowed).
- EventOutcomeDetail should be enclosed in a new EventOutcomeDetailInformation container (with an accompanying <extension> container).
- A new additionalTechnicalCharacteristics container should be included in the DD to accommodate (mainly) format specific technical metadata. It should not be inside the format container.

- A general extension container should NOT be included directly under PREMIS entity containers, except for Rights, as noted above.
- Extensibility may, in future, be extended to cover other semantic units, if demand indicates.
- A new 'Special Topics' section should be included in the DD to explain extensibility, including applicable rules, as above.
- (Additional rules for handling mandatory element requirements, as indicated in the summary (see <http://pec.lib.uchicago.edu:8888/pec/42>) may be included at a later date when needed. At present, all extensible units are optional. The additional rules may just confuse the explanation if included now, so they should be left out.)

2. Next call

Thursday, 8 February, 2008.