PREMIS Editorial Committee Conference Call Notes

16 August 2007

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

1. Group discussion of DD changes

Change 28 - Significant Properties (& Preservation Level)

Issue: Ambiguity in the use of preservationLevel and more guidance and/or structure required on use of significantProperties.

Proposal: Gerard provided further proposals (see http://pec.lib.uchicago.edu:8888/pec/25) based on feedback from the earlier rounds (Feb 2007 [see http://pec.lib.uchicago.edu:8888/pec/16]).

- **Relatedness:** Feedback indicates that preservationLevel & significantProperties are unrelated and should be kept separate. If specific implementations prefer to assign a 'preservationLevel'-type element as part of describing significantProperties, this would be a use outside the current definition in the DD. In such a case, a separate extension unit with a different name could be used in locally for this function.
- Additional Entities and containers to relate these two units seem unwarranted.
- **PreservationLevel:** Make preservationLevel a container, containing a Value, plus a dataConstraint that defines the vocabulary giving meaning to the Value. Include optional units for Rationale (to show the reason that a particular preservationLevel was assigned) and a DateAssigned (to show when).
- **SignificantProperties:** Include some basic structure inside a signifcantProperties container – a Value (not clear what the actual content should be here), a Type (which could note the aspect being described or a particular property – e.g. 'content' or 'structure' etc; 'page size'), and a dataConstraint which defines the vocabulary of Types being used.

Discussion:

- **PreservationLevel:** There was significant debate about the merits of preservationLevel indicating the intended level of preservation or the level implemented.
- 'Intended level' can show what the repository intends to do or is required to do (e.g a customer has paid for this level, or it is mandated as part of a government requirement on archive material). It can also be needed if transferring objects to another repository taking on the responsibility – it shows what level of preservation is expected despite capability.

- 'Intention' may be too 'squishy' – and subject to revision in face of funding or policy changes.
- The only core preservationLevel may be the one that is actually implemented.
- There may be a need for both 'intended' and 'implemented'. In Priscilla's repository, both are needed. Olaf noted that in the archival scene, the intended or required level might be preferable.
- It was suggested that the definition in the DD be changed to make preservationLevel refer only to the level actually implemented. If an intended level is required, this could be added as a separate local unit, of could be added in an additional note field.
- Olaf suggested that a Role unit could be used, to indicate whether the preservationLevel is as implemented or intended. This might also require a controlled vocabulary to define the allowed terms.
- **SignificantProperties:** The proposed structure seems flexible enough to allow additional structure to be defined for significant properties, without locking out any preferred structure that might be revealed through some of the projects under way looking at this issue (e.g. INSPECT). Some further tweaking may be required pending revision of dataConstraint elements.

Decision:

- Further work needed on preservationLevel.
- SignificantProperties seems almost resolved. Tweak once dataConstraint is revised and agreed.

ACTION: Gerard Clifton to revise proposal for preservationLevel, trying combinations of Role, Rationale & Date. Significant Properties can be revised once dataConstraint settled.

New change – Inconsistencies with EventOutcome & EventOutcomeDetail

Issue(s): (see http://listserv.loc.gov/cgi-bin/wa?A2=ind0708&L=premis-ec&T=0&P=661)

1. EventOutcome & EventOutcomeDetail are described in the DD as allowing a text description of an event outcome, code values or error logs, but in the XML schema takes an XML sub-element.
2. EventOutcomeDetail is described in the DD (pg 1-6, description of Event entity) as being repeatable to both a code value and an error log for the event to be recorded, but the DD semantic unit is shown as 'Not Repeatable'.

Proposal (included in the Issue description, by Olaf):

1. Leave it as described in the DD and change the XML-Schema or vice versa
   OR
   handle it by extensibility of the XML-schema.

2. Change the description in the DD [Notetaker's note: not clear whether this means change the description on pg 1-6 which gives the example, or change the semantic unit repeatability to 'Repeatable'.]

Discussion:

- **Issue 1 ('text' vs 'XML'):** This issue may relate to extensibility. When looking at the XML schema in the context of METS the difference between including a text field or an XML function becomes clear. However, it should also be possible to include other XML output (e.g. error logs or other metadata describing the event), so the possibility to extend here is preferred.
- With regard to extensibility in general, there are places where the DD refers to a text description or some additional structure without specifying the structure, so it would seem allowable to include XML structures which offer more flexibility.
- In the DD, it may be better expand on 'text' to include 'or XML'.
- In the DD, it seldom specifies that text fields have to be pure text strings (so the XML 'string' type may not be strictly correct); we may just need to include a new data type that will allow anything to be included here in order to validate. Could change the schema to reflect what the DD says.
- **Issue 2 (repeatability of EventOutcomeDetail):** Options could be to change all the outcome details into one text description and put them all in one EventOutcomeDetail unit, or make the unit repeatable.
- An alternative possibility for content of the EventOutcomeDetail suggested by Priscilla is that if the detail is long (e.g. a long error log), that you might want to instead just point to the metadata from within EventOutcomeDetail, so might want to include a reference rather than the actual detail.
- It seems preferable to allow repeated EventOutcomeDetails to be included in an EventOutcome container, and that the definition of EventOutcomeDetail could be changed to allow text description, references or some other structure to be included – as per the extensibility discussion above.

Decision (not resolved):

- Olaf noted that these offer a solution, but it is not clear that we have made a decision on these points. It would be useful to see an XML schema example to help clarify.
  ACTION: Olaf Brandt and Zhiwu Xie to discuss this offline and come up with an example.

Change 6 – XML Signatures

Issue: May be preferable to use unstructured keyInfo so that established W3C XML Signature elements can be readily used. (See http://listserv.loc.gov/cgi-bin/wa?A2=ind0603&L=pig&T=0&P=53)

Proposal: Make signature elements in line with XML Signatures

Discussion:

- The intention in the DD was to use XML Signatures, but it was cut down to just use keyInfo elements.
- This appears to be just a schema issue – a constraint in the current schema.
- Could also be covered by an extensibility option.

Decision: Remove the internal structure of keyInfo and tweak the schema to align with (or allow use of) XML Signature elements.

2. Other business

Tutorials: No tutorials are planned right now, although there has been an enquiry about holding a tutorial at Dartmouth.

PREMIS contracts: There may be some funding available to undertake small work packages (e.g. guidelines documents, as per the METS user manual). EC members should consider who might be available and can undertake such work

Next call: 30 August.