PREMIS Editorial Committee Conference Call Notes

5 July 2007

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

Apologies: Yaniv Levi, Zhiwu Xie.

1. Update on rights statement proposal

Priscilla has posted two documents on the EC wiki. The original draft proposal (premis rights.doc) was reformatted as a draft Data Dictionary. The draft DD and the examples were sent on June 11, 2007 to John Kunze who asked Kirsten Neilsen at CDL to review, Evan Owens, Hilde van Wijngaarden who asked Barbara Sierman at KB to review, and Karen Coyle. The DD was revised extensively based on comments from Evan, Barbara and Karen to create the June 28 version. The examples document will be revised soon.

ACTION: All PEC members will review the rights documents posted by Priscilla until July 19 and provide feedback to the PEC list. The documents will then be revised if necessary, by Priscilla, and sent to PIG for feedback. PEC will discuss this topic again on August 16th.

2. Making Object Characteristics, Original Name, Storage applicable to representations (continuation of discussion of changes #9, 16, 17)

The discussion from the last meeting continued via email and in the July 5 meeting.

http://listserv.loc.gov/cgi-bin/wa?A2=ind0707&L=premis-ec&T=0&P=462

http://listserv.loc.gov/cgi-bin/wa?A2=ind0707&L=premis-ec&T=0&P=563

Discussion covered the implications of changing some semantic units from "not applicable" for representations to optional (Object Characteristics, Original Name, Storage). Concerns were expressed: that this makes PREMIS less implementation-neutral; that this would reduce or eliminate the distinctions between representation, file and bitstream.

Further discussion clarified that there are at least three distinct types of representations:

1) a single file;

2) complex objects consisting of a group of files that an institution prefers to treat as a unit;

3) representations consisting of a bundled package containing multiple files, e.g. .tar .zip or an XML file bundled in a METS wrapper.

The implications of using PREMIS for databases, and the question of whether databases can be representations will be better understood when some examples can be seen.

ACTION:

Olaf will collect use case scenarios on the wiki, and possibly raise a discussion on the PIG list.

Steve will help to describe the Swedish use case to help clarify the treatment of databases.

All PEC will give more consideration to the implications of opening up the semantic units in question to the representation level.

3. Controlled Vocabularies
Discussion postponed to the next meeting.

4. Revised Data Model
Olaf presented the revised diagrams for the data dictionary which are on the PEC wiki. Changes will be necessary to the textual description of the data model and to the descriptions of some semantic units to correspond with the changes in the diagram.
Recent email discussion proposed generic semantic linkage container for all units and implications on the repeatability of identifers.
http://listserv.loc.gov/cgi-bin/wa?A2=ind0707&L=premis-ec&T=0&P=267

The linking question, and implications on the data model, should be discussed after PEC decides on the construction of linkages.

Units such as linkingAgentRole would not have a home in a generic container.

Repeatability of identifiers -- institutions may require multiple identifiers from multiple schemes. The question was raised as to which identifier would be used in a linkage. There was agreement that for the purposes of PREMIS there needs to be a single core or primary identifier used in linkages. Other secondary identifiers may be recorded in an entity but do not need to be recorded in the linkages with other entities.
It was suggested that this could be done by assigning a role (primary, secondary) to all identifiers. Only the primary identifier would be used in linkages.
ACTION:
Rebecca will clarify the implications of making identifiers repeatable and of adding a role attribute.

5. Event discussions
not discussed

6. Question about translation of PREMIS DD into other languages
Olaf asked Brian about the process of requesting permission to translate the PREMIS DD since it is published with an OCLC copyright statement. Brian says that OCLC asks groups wishing to translate the data dictionary to inform the PREMIS Editorial Committee what they intend to do.

Priscilla suggested creating a Creative Commons License to the data dictionary which would clarify these questions before they arise.

Next call 19 July 2007

Summary of Actions from this call:

ACTION: (Rights)

All PEC members will review the rights documents posted by Priscilla until July 19 and provide feedback to the PEC list. The documents will then be revised if necessary, by Priscilla, and sent to PIG for feedback. PEC will discuss this topic again on August 16th.

ACTION: (Object characteristics for representations)

Olaf will collect use case scenarios for the types of representations on the wiki, and possibly raise a discussion on the PIG list.

Steve will help to describe the Swedish use case to help clarify the treatment of databases.

All PEC will give more consideration to the implications of opening up the semantic units in question to representation level.

ACTION: (Repeatability of identifiers)

Rebecca will clarify the implications of making identifiers repeatable and of adding a role attribute.