PREMIS Implementation Fair Workshop

Angela Dappert
Digital Preservation Coalition
c/o the British Library
96 Euston Road
London, NW1 2DB, UK
angela@dpconline.org

Rebecca Guenther
Library of Congress
101 Independence Ave SE
Washington DC 20540
+1 703-298-0157
rgue@loc.gov

Sébastien Peyrard
National Library of France
Quai François Mauriac
75706 Paris Cedex 13 - France
+33 1 53 79 45 90
sebastien.peyrard@bnf.fr

Robert Sharpe
Tessella
26, The Quadrant,
Abingdon Science Park
Oxfordshire
OX14 3YS UK
Robert.Sharpe@tessella.com

Pauline Sinclair
Tessella
26, The Quadrant,
Abingdon Science Park
Oxfordshire
OX14 3YS UK
Pauline.Sinclair@tessella.com

Eld Zierau
The Royal Library of Denmark
Søren Kierkegaards Plads 1
1219 København K, Denmark
+45 3347 4690
elzi@kb.dk

ABSTRACT
The PREMIS Implementation Fair Workshop is one of a series of events that have been held in conjunction with iPRES in 2009, 2010 and 2012. It is intended to give implementers of the PREMIS Data Dictionary for Preservation Metadata an opportunity to discuss topics of common interest and find out about the latest developments of the standard and of projects using PREMIS.[1] Included in this update will be discussion about forthcoming changes in the PREMIS Data Model and Data Dictionary; the PREMIS OWL ontology; and contributions from implementers and research projects focused on PREMIS and preservation metadata.

General Terms
Documentation, Standardization.

Keywords
Preservation metadata, Preservation Repository implementation, Data Dictionary, Data model.

1. INTRODUCTION
The PREMIS Implementation Fair Workshop is one of a series of events that have been held in conjunction with iPRES. It is intended to give implementers of the PREMIS Data Dictionary for Preservation Metadata an opportunity to discuss topics of common interest and find out about the latest developments of the standard and of projects using PREMIS.

2. OUTLINE OF WORKSHOP CONTENT
2.1 Update on PREMIS activities
A brief overview of PREMIS activities since the last PREMIS Implementation Fair in Oct. 2012 will be given, including changes to the Data Dictionary as part of the development of PREMIS version 3.0. It also includes changes to the Data Model, as detailed below.

2.2 Changes in the PREMIS Data Model
The next major version of the PREMIS Data Dictionary will be released by the end of the year 2013; an update on the main evolutions of the PREMIS Data Model and Data Dictionary will be given. Notably, the revised data model will consider Intellectual Entities differently. Additionally, this new version provides a better way to describe Environments separately from Objects and allow software and hardware registries to be linked to. This modeling work, and the new features that it allows, will be described.

2.3 Changes for preservation policy metadata
The new version of PREMIS also allows the preservation policy applied to preserved digital objects to be recorded in more detail by updating the preservationLevel semantic container.

2.4 PREMIS OWL ontology and id.loc.gov controlled vocabularies
A revised version of the OWL ontology is to be published in Spring 2013, and all the PREMIS controlled vocabularies will be published at the same time. The ontology allows users to express PREMIS using RDF, and to work easily with Linked Open Data, at a time when those technologies are being used more and more in registries (e.g. UDFR, PRONOM). Along with the existing PREMIS XML schema, this provides an alternate PREMIS-endorsed serialization format, and can be leveraged to address problems of distributed preservation metadata across several preservation repositories or format registries and allow PREMIS metadata to be queried easily.

2.5 Implementations
Specific PREMIS implementations will be reviewed and contributions will be solicited from implementers. Specific issues encountered by implementers will be discussed.

2.6 Preservation Health Check report

The Open Planets Foundation and OCLC Research are conducting a pilot that runs through 2013-2014. The activity involves the National Library of France as a pilot site that provides the preservation metadata from their operational repository and deposit systems. The project consists of a quality analysis of the real-life preservation metadata (METS/PREMIS) used by the pilot site, and intends to demonstrate the value of preservation metadata in mitigating risks by aligning the PREMIS Data Dictionary to risk factors. An update will be given on the current state of the project with particular emphasis on the initial outcomes.

3. PREMIS IMPLEMENTATION FAIR: A WORKSHOP SERIES

The PREMIS Implementation Fair at iPRES 2013 will be the fourth in a series. Previous workshops were held at iPRES 2009 in San Francisco, iPRES 2010 in Vienna and iPRES 2012 in Toronto. These events are intended to highlight PREMIS activities, discuss issues concerning implementation, and provide a forum for implementers to discuss their activities, issues, and solutions. Because this is a rapidly changing area, it is important to provide continuous updates.

iPRES2013 and the PREMIS Implementation Fair provides the right framework for engaging the wider digital preservation community in PREMIS activities and fostering the use and implementation of the standard.

4. INTENDED AUDIENCE

The workshop is designed for those involved in selecting, designing, or planning a preservation project or repository using preservation metadata. This includes digital preservation practitioners (digital librarians and archivists, digital curators, repository managers and those with a responsibility for or an interest in preservation workflows and systems) and experts of digital preservation metadata.

Practitioners will be asked to bring a sample of their preservation metadata for the second part of the workshop.

5. SHORT BIOGRAPHIES OF ORGANIZERS

Angela Dappert is a researcher at the Digital Preservation coalition, working on the FP7 project TIMBUS. She also serves on the PREMIS Editorial Committee. In both capacities she is involved with the issues of modeling and defining metadata for computational environments. She has worked at the British Library on data carrier stabilization, digital asset registration, preservation planning and characterization, eJournal ingest, and digital metadata standards. Before this she worked for Schlumberger, the University of California, Stanford University and Siemens. She holds a diploma in Medical Informatics from the University of Heidelberg, an M.Sc. in Computer Sciences from the University of Texas at Austin and a Ph.D. in Digital Preservation from the University of Portsmouth.

Rebecca Guenther is Chair of the PREMIS Editorial Committee, on which she has served since its establishment in 2006. She worked at the Library of Congress on metadata standards in the Network Development Office for 22 years and is currently an independent consultant in New York on metadata development and training; she also continues to work part-time for the Library

of Congress. She was co-chair of the original PREMIS Working Group which developed the *PREMIS Data Dictionary for Preservation Metadata.*

Sébastien Peyrard has worked since 2008 at the National Library of France as a metadata specialist and digital preservation expert in the Models and Standards Office of the Metadata Department (Département de l'information bibliographique et numérique), where he is responsible for maintaining the data model of BnF's preservation repository (SPAR) and its METS and PREMIS implementation. He is a member of the METS Editorial Board and the PREMIS Editorial Committee since 2011.

Robert Sharpe is Tessella's Head of Digital Archiving Solutions. He has worked for Tessella since 1996 and is now responsible for the management of a programme of innovative projects in the archiving and digital preservation domain for a number of customers. This includes the development of Tessella's Safety Deposit Box solution and working on projects for many national archives including the award-winning UK National Archives projects: the Digital Archive, PRONOM and DROID. Robert is also an active participant on a variety of EU-funded research projects and is a member of the PREMIS Editorial Committee**.**

Pauline Sinclair is a consultant in Tessella's Digital Archiving Division. She has guided a variety of prestigious organisations to develop their digital preservation strategy, consider a variety of possible solutions and to define the specification of detailed systems.

Eld Zierau joined the Royal Library of Denmark in 2007, and has since then been involved in digital preservation aspects e.g. preservation metadata, bit preservation, web archiving and participated in the Planets project. She received her computer science degree in 1989, and has worked since then in various industries with similar challenges to long-term preservation today. In 2011 she achieved a PhD degree in digital preservation and in 2013 she joined the PREMIS Editorial Committee.

6. PROCESS FOR SOLICITING CONTRIBUTIONS

Contributions will be solicited from the PREMIS Implementers' Group via its discussion list and invitations will be extended to members of the PREMIS Editorial Committee. The PREMIS Editorial Committee will review all requests. If this workshop proposal is approved, a call will be sent for contributions to the implementation portion and the deadline will be within a month.

7. SPECIAL REQUIREMENTS

A 4h30 half-day workshop is requested to accommodate the variety of presentations that are needed to update the preservation community on preservation metadata implementations, as well as sufficient time to answer the implementers' questions.

8. REFERENCES

[1] PREMIS Maintenance Activity, http://www.loc.gov/standards/premis

[2] PREMIS Editorial Committee. 2012. PREMIS Data Dictionary for Preservation Metadata (Library of Congress). http://www.loc.gov/standards/premis/v2/premis-2-2.pdf.