

Preservación digital buenas prácticas

Eduardo Pablo Giordanino

Sistema de Bibliotecas y de Información

Universidad de Buenos Aires

Facultad de Farmacia y Bioquímica

30 octubre 2014



Preservación digital: buenas prácticas

Temario

- Proyecto PICT-O CIN 0142
- Preservación digital
- Las directrices de la Unesco
- Los metadatos de preservación digital: PREMIS
- Casos de buenas prácticas
 - Biblioteca Nacional de Suecia
 - Biblioteca Real de Dinamarca
 - Centro de Supercomputadoras y Redes de Polonia
- Webgrafía

Proyectos PICT-O CIN

Proyectos de Investigación Científica y Tecnológica Orientados

- Agencia Nacional de Promoción Científica y Tecnológica (ANPCyT)
- Fondo para la Investigación Científica y Tecnológica (FONCyT)
- Consejo Interuniversitario Nacional (CIN)

Proyectos PICT-O CIN

Bonaerense: PICT-O 2010-149

Investigación y desarrollo en repositorios institucionales: aplicaciones y experiencias en universidades de la región bonaerense

Director: Miguel, Sandra (Universidad Nacional de La Plata)

Universidades Participantes: Universidad Nacional de La Plata/Universidad Nacional del Sur/Universidad Nacional de Luján/Universidad Nacional del Centro de la Provincia de Buenos Aires/Universidad Nacional de Mar del Plata/Universidad Tecnológica Nacional – Regional Bahía Blanca

Centro-Este: PICT-O 2010-140

Hacia el desarrollo y utilización de repositorios de acceso abierto para objetos digitales educativos en el contexto de las universidades públicas de la región centro-este de Argentina

Director: San Martín, Patricia (Universidad Nacional de Rosario)

Universidades Participantes: Universidad Nacional de Rosario/Universidad Nacional del Litoral/Universidad Nacional de Córdoba/Universidad Tecnológica Nacional – Regional Santa Fe/Universidad Tecnológica Nacional – Regional Rosario

Centro-Oeste: PICT-O 2010-132

Red COES de repositorios de acceso abierto

Director: García Garino, Carlos Gabriel (Universidad Nacional de Cuyo)

Universidades Participantes: UN La Rioja/UN de Villa María/UN de Córdoba/UN de San Luis/UN de Río Cuarto/UN de Chilecito/UN de San Juan/UN de Cuyo

Proyectos PICT-O CIN

Metropolitano: PICT-O 2010-142

Bases para la puesta en marcha y sustentabilidad de un repositorio digital institucional

Director: Soto, María Araceli (Instituto Universitario Nacional de Arte)

Universidades Participantes: Universidad de Buenos Aires/Universidad Nacional de General Sarmiento/Universidad Nacional de Lanús/Universidad Nacional de Quilmes/Universidad Nacional de Lomas de Zamora/Universidad Tecnológica Nacional – Regional Buenos Aires

NOA: PICT-O 2010-168

Promoción científica, tecnológica y académica en la región NOA mediante repositorios digitales

Director: Etse, José Guillermo (Universidad Nacional de Tucumán)

Sur: PICT-O 2010-139

Diagnóstico y modelado de infraestructura para soporte de repositorios digitales regionales

Director: Mases, Enrique Hugo (Universidad Nacional de Comahue)

Universidades Participantes: Universidad Nacional de La Pampa/Universidad Nacional de La Patagonia Austral/Universidad Nacional de San Juan Bosco/Universidad Nacional de Comahue

Proyecto PICT-O CIN 0142

EJE 3: Estándares, metadatos y software (Eje 3.1) y Preservación digital (Eje 3.2.)

3.2. Preservación digital

3.2.1. Introducción

3.2.2. Los metadatos de preservación digital

3.2.2.1. PREMIS

3.2.2.1.1. Panorama histórico

3.2.2.1.2. El Diccionario de Datos PREMIS

3.2.2.1.3. Razones para usar PREMIS

3.2.2.1.4. Cuestiones de implementación: la conformidad con PREMIS

3.2.2.1.5. PREMIS: situación actual y tendencias

3.2.2.2. METS

3.2.2.2.1. Panorama histórico

3.2.2.2.2. Usos de METS

3.2.2.2.3. Esquema

3.2.2.2.4. Componentes de perfiles METS

3.2.3. Panorama local. Encuesta

3.2.3.1 Introducción

3.2.3.2 Objetivo

3.2.3.3 Metodología

3.2.3.4 Resultados

3.2.3.5 Conclusión

Proyecto PICT-O CIN 0142

3.2. Preservación digital

3.2.4. Panorama internacional

3.2.4.1. La Conferencia Internacional en Preservación de Objetos Digitales (iPRES-2013)

3.2.4.2. Casos relevantes a nivel internacional

3.2.4.2.1. La experiencia alemana: el consorcio Goportis

3.2.4.2.2. La experiencia francesa: la Biblioteca Nacional de Francia

3.2.4.2.3. La experiencia sueca: la Biblioteca Nacional de Suecia

3.2.4.3. La preservación digital en los repositorios de Barcelona

3.2.4.3.1. El Consorcio de Bibliotecas Universitarias de Catalunya

3.2.4.3.2. Estrategia de preservación digital en el CBUC

3.2.4.3.3. Otros casos de preservación digital en Barcelona

3.2.4.3.3.1. Universitat de Barcelona

3.2.4.3.3.2. Universitat Pompeu Fabra

3.2.4.3.3.3. Universitat Politècnica de Catalunya

3.2.5. Aplicación de PREMIS-METS a partir de un caso

3.2.6. Buenas prácticas en preservación digital

3.2.6.1. Las buenas prácticas para preservación digital según la Unesco

3.2.6.2. Las buenas prácticas: iniciativas internacionales

3.2.7. Conclusiones y recomendaciones

3.2.8. Bibliografía

3.2.9. Anexos Eje 3.2. Preservación digital

3.2.9.1. Anexo 1: Declaración de la Unesco sobre Preservación (2012)

3.2.9.2. Anexo 2: Lista de unidades semánticas PREMIS 2.2

3.2.9.3. Anexo 3: Lista de control para uso de PREMIS en un perfil METS

3.2.9.4. Anexo 4. METS. Gráfico y tablas

3.2.9.5. Anexo 5. Modelo de Encuesta

Preservación digital

- Es un conjunto de acciones destinadas a mantener la accesibilidad de los objetos digitales a largo plazo. (*Directrices para la preservación digital*, Unesco, 2003)
- El objetivo de la preservación digital es sostener la capacidad de visualizar, recuperar y utilizar colecciones digitales ante los vertiginosos cambios y obstáculos tecnológicos y organizativos presentados.

Preservación digital : directrices Unesco

- 1) La Gestión de los objetos digitales en la organización
- 2) La planificación del proyecto
- 3) La selección del soporte
- 4) La selección de formatos de archivos y de estándares de datos apropiados
- 5) La validación de los formatos
- 6) Los nombres de los archivos
- 7) La gestión de los archivos en línea
(identificación y resolución permanente)
- 8) La producción de metadatos de buena calidad
(MARC, DCMI)

Preservación digital : directrices Unesco

- 9) La gestión de los archivos (archivos masters almacenados por separado de copias de difusión)
- 10) La seguridad del sistema
- 11) La autenticidad de los objetos
- 12) La formación de recursos humanos
- 13) La protección de la propiedad intelectual
- 14) El mantenimiento del acceso
- 15) La evaluación de los objetos digitales (período de conservación)

Archivo y preservación

Es probable que muchos recursos digitales no puedan ser usados en el futuro.

La información electrónica puede ser alterada, puede devenir inutilizable por la degradación del soporte de almacenamiento, o por el cambio en las tecnologías del hardware o el software.

La migración de formatos y la emulación de programas son algunas estrategias para evitar estos inconvenientes, junto a los metadatos, que aseguran su supervivencia para el futuro.

Cuando se habla de preservar un objeto digital deben registrarse algunos elementos que permitan conocer su la genealogía de un objeto digital (de dónde viene, cómo surgió y cómo cambió), sus características físicas, datos que deben documentarse para poder usarlo con las tecnologías del mañana.

(Giordanino, 2010)

Preservación digital : los metadatos

- En un sistema de preservación digital, la asignación de metadatos a los objetos digitales es una de las primeras medidas para permitir el acceso a los documentos conservados y su gestión y preservación a largo plazo.

Los metadatos de preservación digital

1. Un recurso debe almacenarse de manera segura para que no pueda ser modificado. (*checksum*)
2. Los archivos deben almacenarse en soportes que puedan leer los equipos actuales.
3. Los formatos de archivos más comunes pueden convertirse en obsoletos (migración, emulación)
4. Mantener un registro de las modificaciones de los archivos originales o cambios en la manera de reproducirlos.

(Caplan, 2009)

Los metadatos de preservación digital

- La gestión de la conservación de documentos en soporte electrónico se complica por la rápida obsolescencia de los equipos de reproducción y del software, que son los mayores obstáculos para la preservación digital.
- Los metadatos de preservación son “la información que un repositorio utiliza para llevar a cabo el proceso de preservación digital.”
(PREMIS, 2008a)

PREMIS Data Dictionary

for Preservation Metadata

version 2.2

July 2012

Contents:

Acknowledgments

Introduction

Background

The PREMIS Data Model

General Topics on Structure & Use

Implementation Considerations

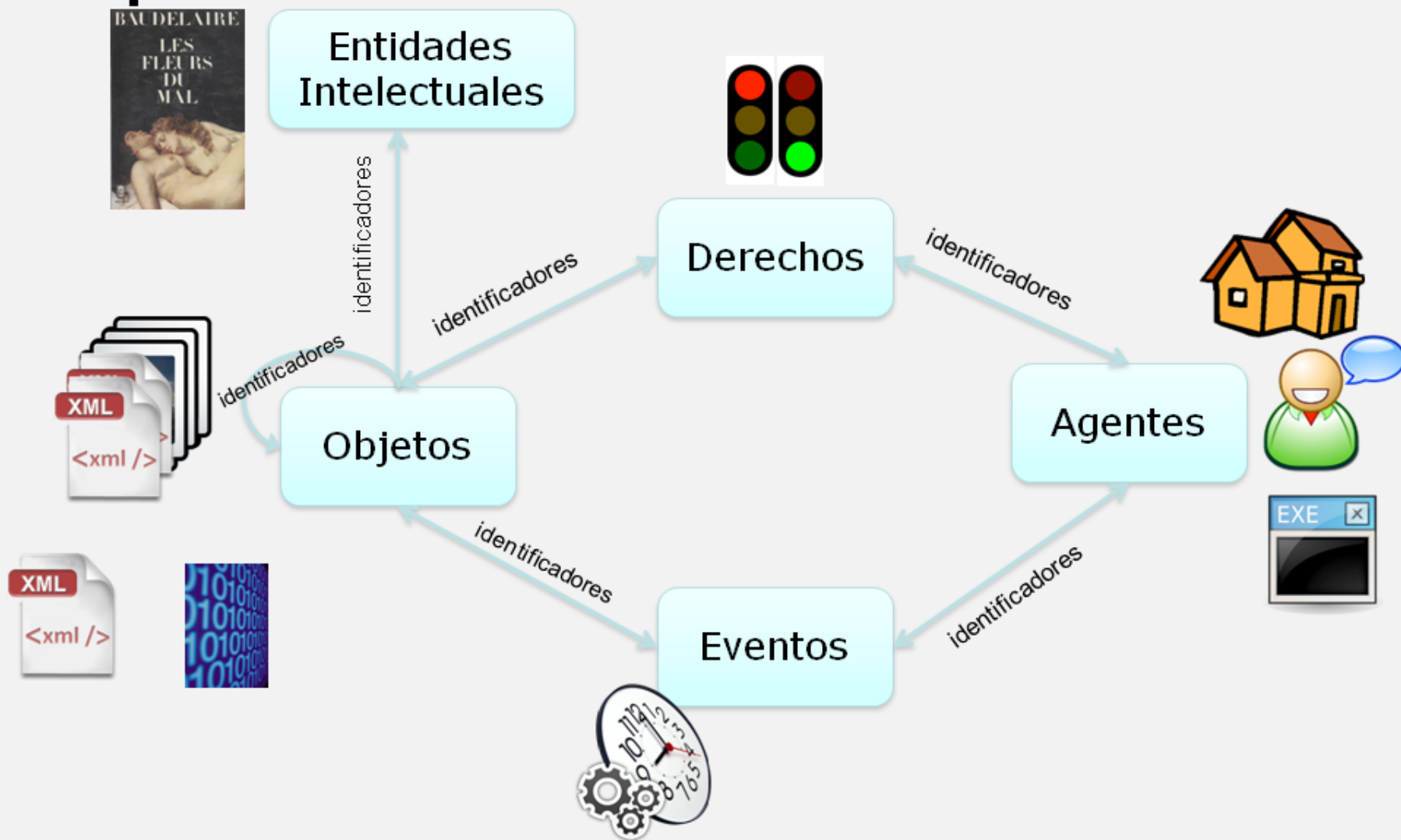
The PREMIS Data Dictionary Version 2.2

Special Topics

Methodology

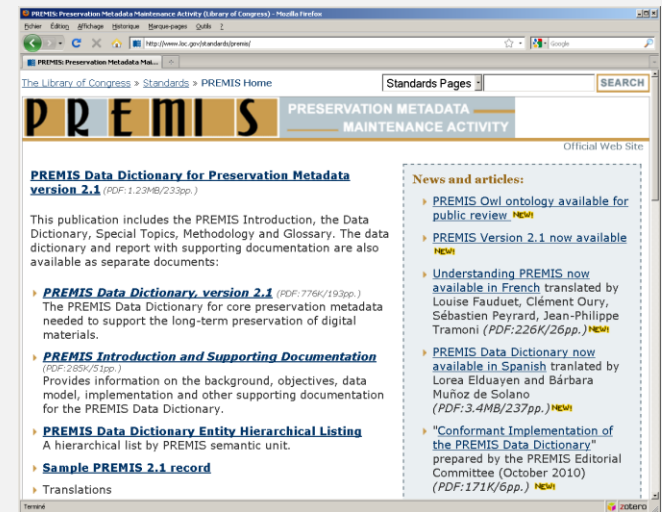
Glossary

Esquema del modelo de datos



Actividad de mantenimiento PREMIS

- Sitio web:
 - hospedado por la *Library of Congress*
 - Información sobre PREMIS, noticias, recursos, eventos
 - Lista de discusión del Grupo de Implementadores de PREMIS (PIG)
- Comité Editorial de PREMIS:
 - Establecen prioridades para el desarrollo de PREMIS
 - Consideran las propuestas de cambios
 - Coordinan la revisión del Diccionario de datos y el esquema XML



<http://www.loc.gov/standards/premis/>

PREMIS in NL Sweden

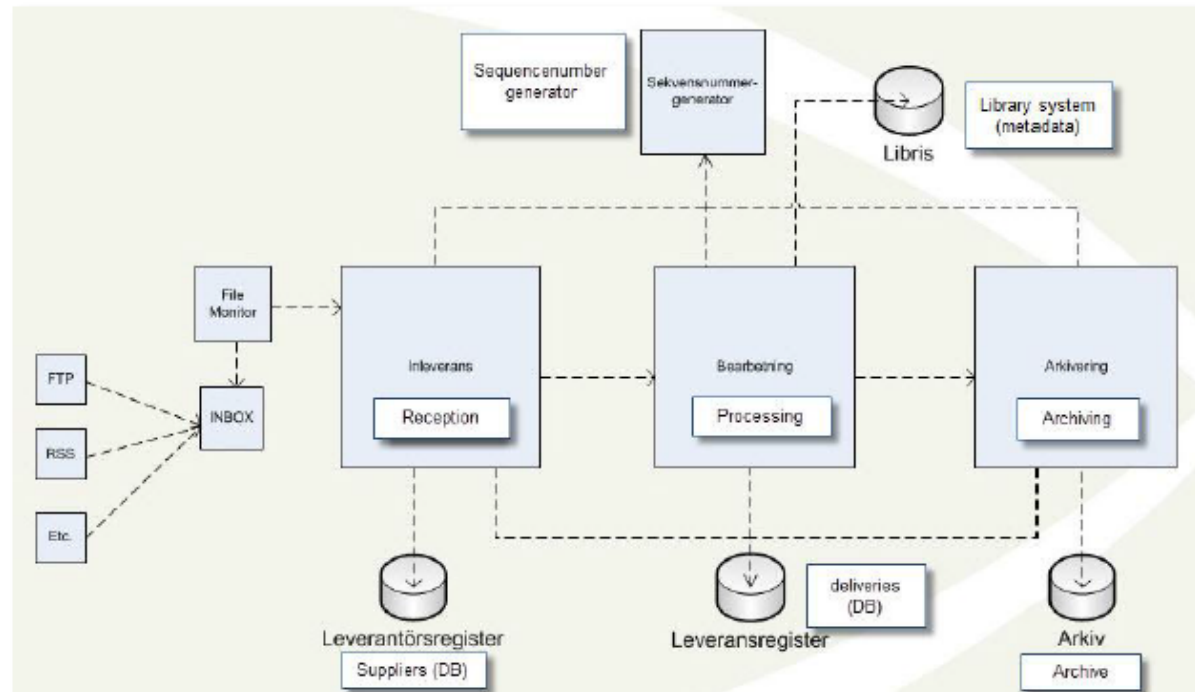
iPRES 2013 –
PREMIS Implementation Fair Workshop

Stina Degerstedt

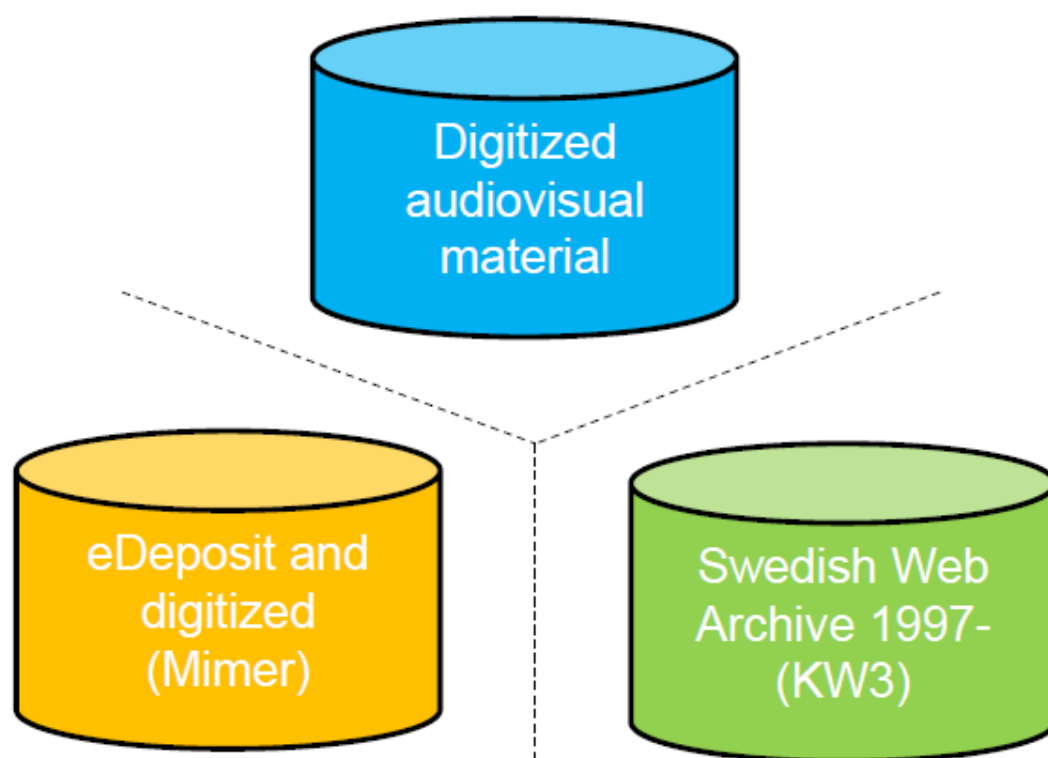


”Mimer” is a platform...

- for managing digital collections at the National Library,
- everything from e-deposit to inhouse digitization
- Conforms to the *OAIS Reference Model*
- Use of well known metadata standards (METS, PREMIS, MODS) – in collaboration with the Swedish National Archives



Not only one digital archive





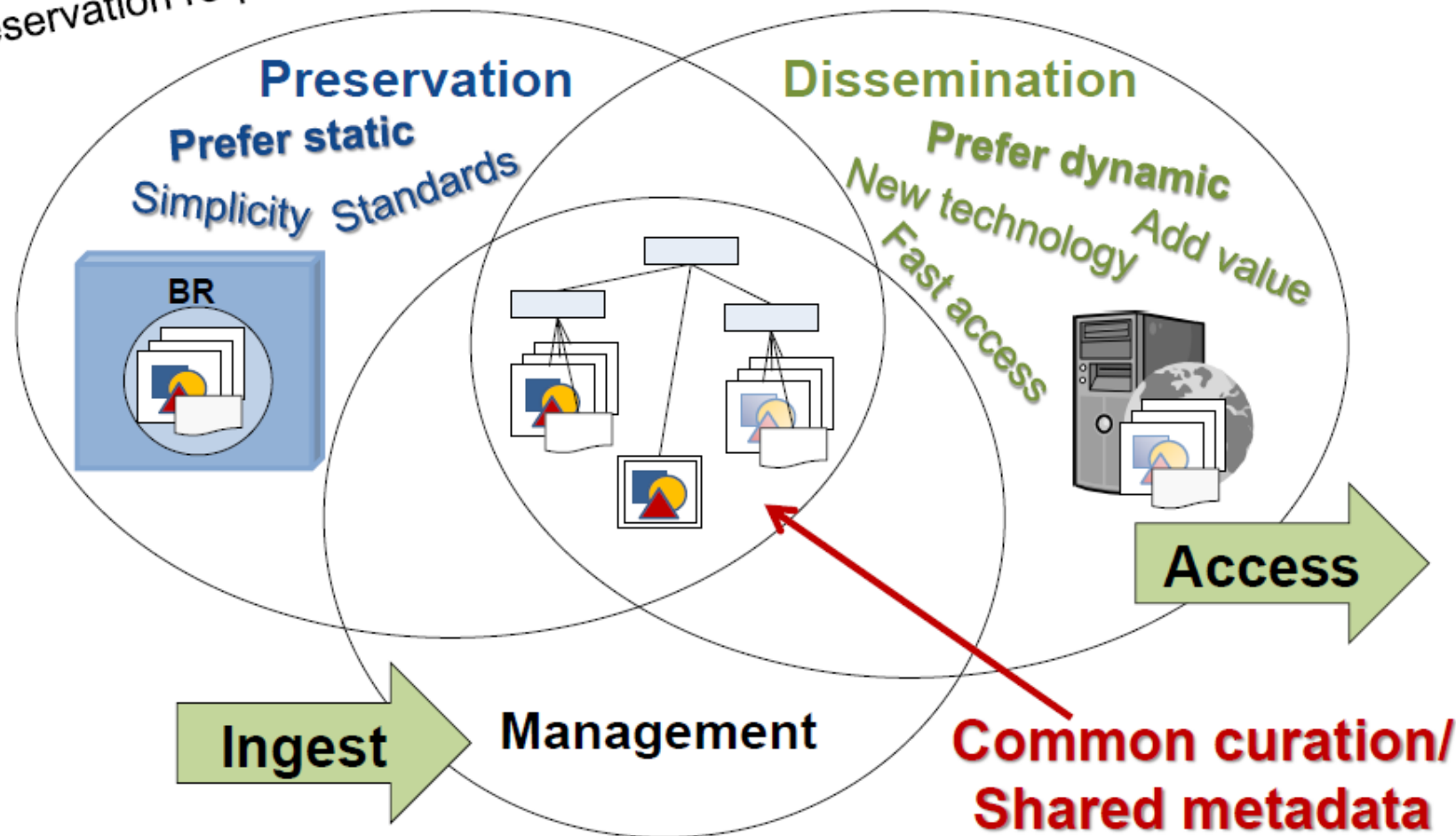
PREMIS Implementation at The Royal Library of Denmark

by Eld Zierau

Digital Library infrastructure



Snapshots
Preservation requires control

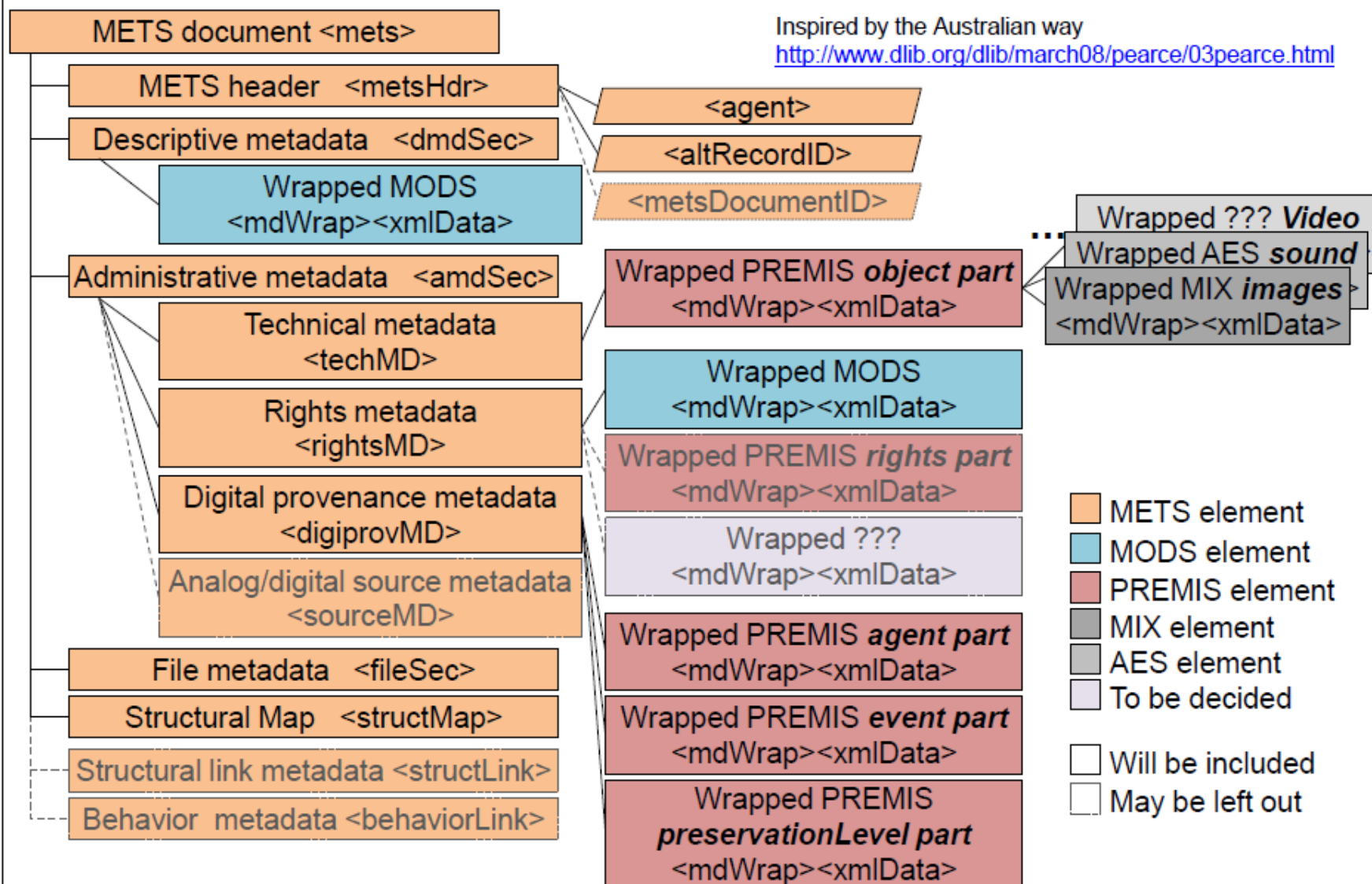


Metadata Standards and use in Preservation



Inspired by the Australian way

<http://www.dlib.org/dlib/march08/pearce/03pearce.html>



Metadata Standards and use in Preservation



example.warc

```
<?xml version="1.0" encoding="UTF-8"?>
<mets xmlns:mets="http://www.loc.gov/METS/" xmlns:premis="info:...">
  <metsHdr CREATEDATE="2013-01-18T19:28:01.025+01:00">
    ...
  </metsHdr>

  <dmdSec CREATED="2013-01-18T19:28:01.035+01:00" ID="Mods1">
    <mdWrap MDTYPE="MODS">
      <xmlData>
        <mods xmlns:xlink="http://www.w3.org/1999/xlink" version="3.4" xsi: ...">
          <genre type="KB Samling">
            F: substitutions-digitaliseret samlings-materiale (som ikke er fortroligt)
          </genre>
          ...
        </mods>
      </xmlData>
    </mdWrap>
  </dmdSec>
```


Metadata Standards and use in Preservation



```
<amdSec>
  <techMD CREATED="2013-01-18T19:28:01.426+01:00" >
    <mdWrap MDTYPE="PREMIS:OBJECT">
      <xmlData>
        <object xmlns:xlink="http://www.w3.org.1999/xlink" xsi: ...">
          <objectIdentifier>
            <objectIdentifierType>UUID</objectIdentifierType>
            <objectIdentifierValue>
              54d153d0-0099-11b2-9397-00505645645
            </objectIdentifierValue>
          </objectIdentifier>
          <significantProperties>
            <significantPropertiesExtension>
              <mix xsi:schemaLocation="http://www.loc.gov/mix/...">
                <BasicDigitalObjectInformation>
                  ...
                </mix>
              ...
            </significantPropertiesExtension>
          </significantProperties>
        </object>
      </xmlData>
    </mdWrap>
  </techMD>
</amdSec>
```

POZNAŃ **SUPERCOMPUTING** AND **NETWORKING** **CENTER**



PREMIS and METS: dArceo case study

Tomasz Parkoła (tparkola@man.poznan.pl)
Network Services Department, PSNC

Digital libraries infrastructure in the PIONIER network

- Developed since 2002
- Consists of around 100 digital libraries
- Collaborative effort of several hundreds of research and cultural institutions including local networking and computing centers
- Over 1.7M digital objects available on-line for the information society



Toolset for Polish digital libraries

Developed and provided by PSNC

tools



projects



Webgrafía

- Biblioteca de Catalunya, Col·legi Oficial de Bibliotecaris-Documentalistes de Catalunya, Consorci de Biblioteques Universitàries de Catalunya (CBUC). (2013). *Estàndards de digitalització: requeriments mínims*. Barcelona. <http://hdl.handle.net/2072/212855>
- Biblioteca de la Universidad de Cornell (2005). *Llevando la teoría a la práctica: tutorial de digitalización*. <http://www.library.cornell.edu/preservation/tutorial-spanish/contents.html>
- Borbinha, J., Michael, N. & Knight, S. (Eds.). (2013). *Proceedings of the 10th International Conference on Preservation of Digital Objects, Lisboa, 2013*. Lisboa: Biblioteca Nacional de Portugal. <http://purl.pt/24107>
- Degerstedt, S. (2013). National Library of Sweden's implementation of PREMIS. In PREMIS. *PREMIS Implementation Fair at iPres 2013 in Lisbon, Portugal: Agenda*. <http://www.loc.gov/standards/premis/premis-implementation-fair-agenda-2013.html>
- Digital Preservation Coalition (2008). *Digital Preservation Handbook*. Recuperado el 17 de septiembre de 2013, de <http://www.dpconline.org/vendor-reports/download-document/299-digital-preservation-handbook.html>
- Fauduet, L. & Peyrard, S. (2013). Destination: Shared Repository. The National Library of France's Journey to Third-Party Archiving. In Borbinha, J., Michael, N. & Knight, S. (Eds.). *Proceedings of the 10th International Conference on Preservation of Digital Objects, Lisboa, 2013*. (pp. 172-176). Lisboa: Biblioteca Nacional de Portugal. <http://purl.pt/24107>
- Gallart, N. & Keefer, A. (2007). *Preservació de dipòsits digitals*. <http://www.recercat.net/bitstream/2072/4078/1/0703Preservacio.pdf>
- Lindlar, M., Friese, Y., Müller, E., Bähr, T. & Trosdorf, A. (2013). Benefits of geographical, organizational and collection factors in digital preservation cooperations: The experience of the Goportis consortium. In Borbinha, J.; Michael, N. & Knight, S. (Eds.). *Proceedings of the 10th International Conference on Preservation of Digital Objects, Lisboa, 2013*. (pp. 110-117). Lisboa: Biblioteca Nacional de Portugal. <http://purl.pt/24107>
- Parkoła, T. (2013). PREMIS usage in PSNC-developed dArceo long-term preservation services. In PREMIS. *PREMIS Implementation Fair at iPres 2013 in Lisbon, Portugal: Agenda*. <http://www.loc.gov/standards/premis/premis-implementation-fair-agenda-2013.html>
- PICTO-CIN 2-0142 (2014). *Bases para la puesta en marcha y Sustentabilidad de un Repositorio Digital Institucional*. María Araceli Soto (coord.). Buenos Aires, Instituto Universitario Nacional del Arte. Crítica de Artes, 2014.
- PREMIS (2013). PREMIS Implementation Fair at iPres 2013 in Lisbon, Portugal: Agenda. <http://www.loc.gov/standards/premis/premis-implementation-fair-agenda-2013.html>
- Unesco (2012). *UNESCO/UBC Declaración de Vancouver: La Memoria del Mundo en la era digital: digitalización y preservación*. http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CI/CI/pdf/mow/unesco_ubc_vancouver_declaration_es.pdf
- Universidad de Barcelona (2011). *Política de acceso abierto en la Universidad de Barcelona. Aprobado en Consejo de Gobierno de 7 de junio de 2011*. http://diposit.ub.edu/dspace/bitstream/2445/27710/1/2011_06_UB_OA_politica_spa.pdf
- Universitat Pompeu Fabra. Biblioteca y TIC · CRAI (2013). *Misión*. <http://www.upf.edu/bibtic/es/coneixer/missio.html>
- Witten, I.H., Bainbridge, D. & Nichols, D.M. (2010). *How to Build a Digital Library*. 2nd ed. Burlington, MA: Morgan Kaufmann Publishers.

¿Preguntas?

¡Gracias por su atención!

egiordanino@sisbi.uba.ar

<http://twitter.com/egiordanino>

