Towards Interoperable Preservation Repositories

TIPR

DLF Spring Forum, 2009
Joseph Pawletko (NYU), Priscilla Caplan (FCLA), Bill Kehoe (CUL)
Presentation Overview

- Motivation & Goals
- Analysis
- TIPR Exchange Package (TEP)
- Transfer Tests
- Progress to Date / Next Steps
- Questions
## Motivation:

<table>
<thead>
<tr>
<th>Repository Requirement:</th>
<th>Mitigation Strategy:</th>
</tr>
</thead>
<tbody>
<tr>
<td>must not lose data</td>
<td>geographically distribute copies, ideally in heterogeneous systems</td>
</tr>
<tr>
<td>accommodate evolving technologies</td>
<td>upgrade/replace repository implementation</td>
</tr>
<tr>
<td>succession plan</td>
<td>be able to transfer AIPs to another institution</td>
</tr>
</tbody>
</table>

All strategies benefit from AIP Transfer.
TIPR Goals:

- Demonstrate feasibility of repository-to-repository transfer of rich AIPs
- Identify and resolve issues that impede such transfers
- Develop a usable, standards-based transfer format
- Disseminate results to the international preservation community and the relevant standards activities
TIPR Analysis Phase:

- Analyze Repository and AIP structures to determine transfer format
- Determine Repository modifications required to Ingest and Export Exchange Packages
TIPR Partners

- FCLA Archive
  - producer
  - producer

- TIPR Exchange Format

- Cornell Archive
  - producer
  - producer
  - producer

- NYU Archive
  - producer
  - producer
  - producer
Florida Center for Library Automation (FCLA)

- runs the **DAITSS** Preservation Repository

**Team**: Priscilla Caplan (PI), Franco Lazzarino, Randy Fischer, Marly Wilson
Cornell University Libraries (CUL)
- runs an aDORe-based repository
- migrating to Fedora
- Team: Oya Rieger, Bill Kehoe, Rick Silterra
New York University Libraries (NYU)
- runs a DSpace-based repository
- **Team**: Dr. Michael Stoller, Joe Pawletko, Rasan Rasch
Terms as defined by PREMIS

- Intellectual Entity:
  “a set of content that is considered a single intellectual unit for purposes of management and description”.

- Representation:
  “the set of files, including structural metadata, needed for a complete and reasonable rendition of an Intellectual Entity.”
• **One AIP per Intellectual Entity**

• **Retain First and Latest Representations**

• **Discard Intermediate Representations**
TIPR
AIP Structures
CUL

- **One AIP per Representation**
- **Retain All Representation AIPs (including “intermediates”)**

![Diagram of Intellectual Entity]

- Representation $R_0$
- Representation $R_1$
- Representation $R_2$
- Representation $R_n$
• One AIP per Intellectual Entity
• Retain all Representations (including “intermediates”)

Intellectual Entity

Representation $R_0$

Representation $R_1$

Representation $R_2$

Representation $R_n$
## Repository / AIP Heterogeneity:

<table>
<thead>
<tr>
<th>Repository Implementation</th>
<th>FCLLA</th>
<th>CUL</th>
<th>NYU</th>
</tr>
</thead>
<tbody>
<tr>
<td>Repository Implementation</td>
<td>DAITSS</td>
<td>aDORe -&gt; Fedora</td>
<td>DSpace-based</td>
</tr>
<tr>
<td># of AIPs per n representations</td>
<td>1</td>
<td>n</td>
<td>1</td>
</tr>
<tr>
<td>retain intermediate representations?</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
</tr>
</tbody>
</table>
TIPR Exchange Package (TEP)

- TEP design goals:
  - be able to accommodate different AIP structures
  - contain sufficient data for receiving repository at:
    - TEP level
    - Representation level
  - use standards: METS and PREMIS
TEP: Minimal Structure

./tipr.xml

./tipr-digiprov.xml

./tipr-rights.xml

./tipr-rep-1.xml

./tipr-rep-1-digiprov.xml

./files/
TEP: Minimal Structure

./tipr.xml

./tipr-digiprov.xml

./tipr-rights.xml

./tipr-rep-1.xml

./tipr-rep-1-digiprov.xml

./files/
TEP: Minimal Structure

Representations-level information
TEP: Minimal Structure

METS document describing Sender. Uses mdRefs to point to rights, digiprov, and representation files.
TEP: Minimal Structure

./tipr.xml

./tipr-digiprov.xml

./tipr-rights.xml

./tipr-rep-1.xml

./tipr-rep-1-digiprov.xml

./files/

PREMIS document containing TEP-level digiprov information
TEP: Minimal Structure

./tipr.xml

./tipr-digiprov.xml

./tipr-rights.xml

./tipr-rep-1.xml

./tipr-rep-1-digiprov.xml

./files/

PREMIS document containing TEP-level rights information
TEP: Minimal Structure

./tipr.xml
./tipr-digiprov.xml
./tipr-rights.xml
./tipr-rep-1.xml
./tipr-rep-1-digiprov.xml
./files/

METS document describing representation 1
TEP: Minimal Structure

./tipr.xml

./tipr-digiprov.xml

./tipr-rights.xml

./tipr-rep-1.xml

./tipr-rep-1-digiprov.xml

/files/

PREMIS document containing representation 1 digiprov information
TEP: Minimal Structure

- ./tipr.xml
- ./tipr-digiprov.xml
- ./tipr-rights.xml
- ./tipr-rep-1.xml
- ./tipr-rep-1-digiprov.xml
- ./files/

directory containing the representation files
TEP: Optional Files

```/tipr.xml.sig```

a stand-alone digital signature in OpenPGP format generated using sender’s private key, and `tipr.xml`
Each `tipr-rep-n.xml` must be accompanied by a corresponding `tipr-rep-n-digiprov.xml`. Information for additional Representations.
Transfer Tests

- Will run two different transfer scenarios
  - Broadcast Transfer
  - Ring Transfer

- Will analyze test results against expectations
  - will use results to improve TEP structure
FCLA sends $\text{TEP}_{\text{FCLA}}$ to CUL and NYU. CUL and NYU Ingest $\text{TEP}_{\text{FCLA}}$. 
CUL sends $\text{TEP}_{\text{cul}}$ to FCLA and NYU. FCLA and NYU Ingest $\text{TEP}_{\text{cul}}$. 
NYU sends $\text{TEP}_{\text{NYU}}$ to FCLA and CUL. FCLA and CUL Ingest $\text{TEP}_{\text{NYU}}$. 
FCLA sends TEP\textsubscript{FCLA} to CUL. CUL Ingests TEP\textsubscript{FCLA}.
CUL sends $\text{TEP}_{\text{FCLA, CUL}}$ to NYU.
NYU Ingests $\text{TEP}_{\text{FCLA, CUL}}$. 
NYU sends $\text{TEP}_\text{FCLA,CUL,NYU}$ to FCLA.
FCLA Ingests $\text{TEP}_\text{FCLA,CUL,NYU}$. 
Progress to Date

- Defined TEP Structure
- Each Partner created TEPs for sample AIPs
- TIPR Specification draft on Partner site
- Developed Schematron tests for TEP METS, PREMIS
- Coding in progress
  - FCLA can programmatically generate TEPs
  - CUL and NYU development in progress
Issues Under Investigation:

- **Rights:**
  - What rights pertain to an AIP when transmitted from one repository to another?
  - Can sending repository restrict what receiving repository can do with TEP?
  - Seeking clarification from legal experts

- **Significant Properties**
  - how to communicate via TEP?
Next Steps:

- Finish coding
- Run tests
- Analyze results
- Refine transfer format
- Communicate findings
Acknowledgements

- TIPR is funded by a generous grant from the Institute of Museum and Library Services (IMLS)
Questions?

- Thank you for your time...
TIPR: Interchange you can believe in!

http://wiki.fcla.edu:8000/TIPR/