Selecting an Electronic Records Repository Platform

How we conjured something from nothing

A Presentation By
Bryan Collars and Brian Thomas
South Carolina Department of Archives and History
BPE 2015
Topics

• A Brief Background of Electronic Record Programs at the SCDAH

• Evaluation of Repository Platforms

• Demonstration of VR
But First
Background

• When I entered the field electronic records generally induced......
History
• PeDALS
• Then 2008 happened…
Background

- Theory but very little practice (Camp Pitt?)
- Starts and Stops and Re-Starts and Communes
- NDIPP
- PeDALS
- Then in 2008

6 long years later........
2014: A new attempt

• New Director (2009)
• New hardware funding  
  – Installed August 2015
• Search for an Electronic Records Archivist
• June 2014 – New Electronic Records Archivist hired
• Process begins, 4 steps
The plan

- Research/literature review
- Policy writing (draft)
- Testing and evaluation
  - Self assessment
  - Testing
- Final recommendation and acquisition
Research/literature review

OAIS Standards

- CCSDS 652.0-M-1: Audit and Certification of Trustworthy Digital Repositories
- CCSDS 650.0-M-2: Reference Model for an Open Archival Information System
- CCSDS 651.0-M-1: Producer-Archive Interface ... Standard
- LOC: TRAC Criteria and Checklist
  – Most useful of the bunch

Most useful of the bunch
Research/literature review

Metadata Standards
- PREMIS
- Dublin Core
- EAD
- EAC-CPF
- ISAAD(G)
- ISAAR
- PROV
- Linked Data
- METS
- MODS
More Research

- Articles
- Format standards/recommendations
- Repository platform options
- NOT POWRR (why)
• Policy writing
  – Why policy writing before testing?
  – TRAC criteria used guidelines
  – Metadata requirements drafted
Self-assessment

• Self-assessment
  – What are the current financial resources?
  – What are the future financial resources?
  – Technology resources?
  – What about staffing?
  – What about cost of maintenance?
## Policies and self-assessment

<table>
<thead>
<tr>
<th>Function</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Submission*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Manual Submission option</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Web-based submission option</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Batch Submission option</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Submission formats</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Costs</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Open-source v. proprietary</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subscription costs for maintenance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mandatory subscription</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Installation costs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Can maintain if no longer able to pay</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Installed by developer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Installed by 3rd party</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Longevity in the field</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Administration</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Can create workflows</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Can create permissions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Allows virus scan</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Allows quarantine</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Can create thumbnails/access copies</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Format monitoring</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Back-up</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normalization</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Versioning</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Import anything</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plug-ins by company</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plug-ins by 3rd party</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plug-ins for a fee</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plug-ins for free</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Custom API/submission interface</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Easy set-up (1-10)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Easy maintenance (1-10)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data types</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Metadata</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>METS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MODS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PREMIS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EAD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Access metadata</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Structural metadata</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preservation metadata</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Content management/access</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Double as a content management system</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inter-operable with a content management system</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OAIPMH</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Text indexing and searchable</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frequency in use</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linux base</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Windows base</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Submission*

Manual Submission option

Web-based submission option

Batch Submission option

Administration

Can create workflows

Can create permissions

Allows virus scan

Metadata

Other

Frequency in use

Linux base

Windows base

Text indexing and searchable

Access metadata

Structural metadata

Preservation metadata

Plug-ins for a fee

Plug-ins for free

Custom API/submission interface

Easy set-up (1-10)

Easy maintenance (1-10)

Database types
Final Criteria reached

Tier 1
- Integrated Repository Platform
- Automated repository services
- Integrated access tools

Tier 2
- Unintegrated Repository Platform
- Some or all automated repository services
- Manual DIP connection to access or mediated access

Tier 3
- Manual preservation tools resulting in a repository
Testing

Tier 2
• DAITSS
• Digital Preservation Software Platform (DPSP)
• Fedora

Tier 1
• Hydra (not tested)
• Dspace (previous experience)
• Islandora
• Archivematica/AtoM
• Preservica

Tier 3 tools tested but not part of this presentation
Method: Virtual Machines

**Prep**
- Read platform documentation
- Read dependency documentation

**Install**
- VMWare or VirtualBox
  - Install OS (typically Ubuntu 12.04)
  - Lesson: follow direction to the letter

**Test**
- Checklist of functionality/wishlist
- Usability
- Side-by-side testing
Final Recommendation

Preservica