Select for Success: Key Principles in Assessing Repository Models

Oya Y. Rieger
Cornell University Library
April 2007

Within a lifecycle framework, digital curation involves a series of technical, intellectual, and managerial activities in support of stewardship for digitized or born-digital information assets.
What is a repository system?

Capture, store, index, manage, preserve, and deliver digital objects.

- Digital asset management
- Discovery & retrieval
- Institutional visibility
- Academic support
- Scholarly communication
- Business procedures support
- Innovation in information management

DSpace
Fedora
GreenStone
ContentDM
Sakai
DiVA
ePrints
Digital Commons
DPubS
ContentDM
Open Journal System
aDORe
Factors in Choosing a Repository Model

<table>
<thead>
<tr>
<th>Development characteristics</th>
<th>Support of archival business requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Financial sustainability</td>
<td>Content type characteristics</td>
</tr>
<tr>
<td>Digital library infrastructure</td>
<td>Preservation functionality</td>
</tr>
<tr>
<td>Interoperability and support for standards</td>
<td>Usability (staff and end-user)</td>
</tr>
<tr>
<td>Institutional policies and practices</td>
<td>Search, browse, access features</td>
</tr>
</tbody>
</table>


**Research Libraries Group and OCLC. Trustworthy Repositories Audit & Certification**
Key Principles in Selecting a Repository Model

1. Identify Key Stakeholders
How about OAIS?

What is your business case?
Let’s implement Fedora

Focus on intellectual content
Why would I give you my data?

- Build awareness
- Gather feedback
- Build trust
- Get support
- Expand resources
- Understand risks

Benefits of Stakeholder Analysis
2. Conduct Needs Assessment

- Document type
- Condition
- Metadata attributes
- Selection criteria
- Usage restrictions
- Relation to other collections

- Services
- Use type
- Frequency of use
- Use mode
- User support needs

- Staff and skills
- Systems, hardware, software
- Stakeholders
- Organizational guidelines and policies
- Master plans and strategies

3. Explore Resource Requirement

- Start-up costs to be in a range of $8,000-$1,800,000 (with a mean of $182,550)
- An average ongoing operating cost of $113,500


LIFE (Life Cycle Information for E-Literature)
University College London (UCL), the British Library, JISC
Case Study: arXiv

Staff = Leadership, management, support, tech development, admin support (9 staff – 4 FTE)

Equipment = Servers, back-ups, amortization, storage costs, maintenance

Other = Meetings, supplies, communication, software, etc.
4. Understand the Existing Human Landscape

- Equally important is taking into consideration the existing and evolving work culture and practices.
- Traditions and work practices of different communities shape their acceptance and use of technologies.
  - Relevant social groups (stakeholders)
  - Interpretive flexibility
  - Appropriation

Approaches that try to “move” faculty and their deeply embedded value systems directly toward new forms of archival systems are destined to fail.

Key Principles in Selecting a Repository Model

- Identify key stakeholders
- Conduct needs assessment analysis
- Explore resource requirements
- Understand the existing human landscape

Conclusions

- Flexible and scalable repositories

At Johns Hopkins, we are promoting the idea that applications should access repositories through an abstract, repository agnostic layer, rather than through custom application to repository integrations.

Choudhury and Martino, 2005
Conclusions

- Flexible and scalable systems
- Web services

Conclusions

- Flexible, scalable
- Web services
- Repurposing
Conclusions

- Flexible and scalable
- Web services
- Repurposing
- New information chain

Van de Sompel et al, D-Lib Magazine, September 2004