Cooperative Strategies for Distributed Digital Preservation

Toward a Reference Model for Distributed Digital Preservation

Katherine Skinner, Executive Director, Educopia Institute
Martin Halbert, Dean of Libraries, University of North Texas

CNI 2010 Spring Forum, Baltimore MD
April 13, 2010
Presentation Overview

- Educopia Institute overview
- Field formation issues
- Notes from the DDP field on Organizational Strategies
- Collaborative Preservation Models
- Need for a DDP Reference Model

Skinner & Halbert, 2010
Educopia Institute

- 501(c)3 founded in 2006 to foster collaboration between cultural memory organizations in the production, dissemination, and preservation of digital resources
- A de-centered, catalytic change agent
- Common (neutral) ground for academic and research-oriented support

Skinner & Halbert, 2010
Educopia Programs

- MetaArchive Cooperative
- InterState Archive (forthcoming)
- Scholarly communication program (forthcoming)
- Educopia Consulting
MetaArchive Cooperative

- Established in 2004 (support from NDIIPP and NHPRC), preserving content for 15 members
- Uses LOCKSS software to provide peer-to-peer distributed digital preservation infrastructure
- Sustainable organizational framework: Membership organization with a 501c3 host
- 254 TB network capacity (and growing)
- Compliant as a Trustworthy Digital Repository (2009 TRAC audit available on our site)
1. Cultural memory organizations must maintain their historical role as cultural stewards
   ▪ Preservation of digital assets as corollary to preserving physical ones
   ▪ Need in house expertise and knowledge
   ▪ Value of curators and librarians and archivists

Skinner & Halbert, 2010
2. Capitalize on cultural memory organizations’ proven preservation methodologies
   ▪ Replication of content
   ▪ Distribution of content
   ▪ Partnering to keep costing affordable
Current Members
Auburn University
Boston College
Clemson University
Florida State University
Folger Shakespeare Library
Georgia Tech
Library of Congress
Penn State University
PUC Rio de Janeiro
Rice University
University of Hull
University of Louisville
University of North Texas
University of South Carolina
Virginia Tech

Current Affiliates
Library of Congress
NDLTD
SDSC Chronopolis
InterState Archive (forthcoming)

- Hub-and-spoke architecture for digital preservation
- Developing organizational apparatus
  - How do hub and spokes interact?
  - How long must hubs commit to preservation efforts?
  - How do hubs interact w/ each other?
  - What policies can govern community behavior and what areas can hubs define for themselves?

Skinner & Halbert, 2010
Educopia Consulting

- Consulting
  - policy creation, business planning, PLN hosting, TRAC audit preparation
- Publishing
- Training
- Event planning (Aligning National Digital Preservation Strategies Conference 2011)
- Research

Skinner & Halbert, 2010
Field formation 101

Once prevalent thought:
- Business logics and models exert dangerously stable control over production, distribution, and reception of cultural products
  (Adorno, Bagdikian, Hall, Hebdige)

Now, more common understanding:
- Fields/genres are social constructions with organizing principles that are always in flux
  (DiMaggio, Leblebici, Peterson, Negus)

Skinner & Halbert, 2010
Field formation 101

- Cultural processes of production, distribution, and reception depend upon networks of people

- Significant cultural events, including activism, often allow specifically for the emergence of new fields

- Watch the fringes of the field for key entrepreneurs
  - broadcast music
  - Jukeboxes

Skinner & Halbert, 2010
Field formation, 101

- Print to digital – moment of transformative social change

- Our changing modes of communication are changing scholarship
  - production, dissemination, reception, & preservation

- We play a pivotal role—whether passively or actively—in how this field is emerging

Skinner & Halbert, 2010
Field formation studies help us...

- Recognize what’s at stake
  - Ultimately, our own institutions
    - Missions—access and preservation
    - Outsourcing a core mission is a dangerous proposition

- Understand how to structure our activities
  - Use standards wisely
  - Think models, not mandates
  - Breakthrough-based coordinated efforts

Skinner & Halbert, 2010
Observations from the Field

- What organizational strategies work, and what strategies don’t work?
- What lessons can we learn from the successes thus far?
- What innovations are still needed?
Coordinated efforts with clear leadership and shared responsibility/governance

Distributed collaboratives

Standards used as a touchstone for innovation

Development of protocols

Overextended lead institution trying to gain attention from (or wrestle with!) partner-peers

Centralized approaches

Standards used as inflexible rules to which one must conform

Development of tools

Skinner & Halbert, 2010
<table>
<thead>
<tr>
<th>Less effective Thinking</th>
<th>More Effective Thinking</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sustainability via building the best product</td>
<td>Sustainability via building the best relationships</td>
</tr>
<tr>
<td>Someone else will solve the problem</td>
<td>Everyone needs to engage on this issue</td>
</tr>
<tr>
<td>Preservation and access must be together to be economically feasible</td>
<td>Preservation must be given visibility and meaning for content producers, users, and owners</td>
</tr>
</tbody>
</table>

Skinner & Halbert, 2010
Bottom line:

BUILD TOWARD SUSTAINABLE STRATEGIES, NOT SUSTAINABLE PRODUCTS

Skinner & Halbert, 2010
Gap Analysis of What the Field Lacks Right Now

- Analysis / abstracted model for distributed digital preservation:
  - Peer-to-peer roles
  - Hub-and-spoke styled preservation relationships
  - Centrally orchestrated
- Ingestion pathways
- Contingent elements

Skinner & Halbert, 2010
OAIS Reference Model

**Diagram:**

- **Preservation Planning**
- **Data Management**
- **Archival Storage**
- **Access**
- **Administration**

**Processes:**

- **Ingest**
- **Descriptive Info**
- **SIP**
- **AIP**
- **DIP**

**Data Flows:**

- **PRODUCER**
- **CONSUMER**
- Queries result sets orders

**Source:**

Skinner & Halbert, 2010
Roles in the Stewardship Network

Committed Content Custodians

Communities of Practice and Information Exchange

Services

Capacity Building

Source: “Since we met last year...” Plenary, Martha Anderson, National Digital Information Infrastructure and Preservation Program Annual Partners Meeting 2008
We need a reference model similar to the OAIS, but for Distributed Digital Preservation alliances, systems, and strategies.

Such a reference model should abstract the functions and logical potential relationships between entities seeking to work together to further digital preservation aims.

This model should inform collaborative efforts between different groups seeking to preserve digital information.

It should also provide a common vocabulary for interoperable systems development.

Skinner & Halbert, 2010
A cluster of DDP organizations including MetaArchive, Chronopolis, and others are considering a collaborative project to develop such a DDP Reference Model.

Are in conversations with LC and other agencies about hosting an initial planning meeting to study this issue.

If you are interested, please contact us.
Contact

Martin Halbert
martin.halbert@unt.edu

Katherine Skinner
katherine.skinner@metaarchive.org