

The MetaArchive of Southern Digital Culture

Martin Halbert
Emory University
DLF Spring Forum
San Diego, California
Thursday, April 14, 2005



MetaArchive Project Summary



Six partner institutions will collaborate with LoC on a three year \$1.3M effort to develop a cooperative for the preservation of digital content with a particular content focus:

Southern culture and history

Project Goals



1. Create a conspectus of digital content within the subject domain held by the partner sites
2. Harvested body of the most critical content to be preserved (3 terabytes, w/ capability to expand)
3. Develop a model cooperative agreement for ongoing collaboration and sustainability
4. Distributed preservation network infrastructure based on the LOCKSS software

Key Features of a Secure MetaArchive



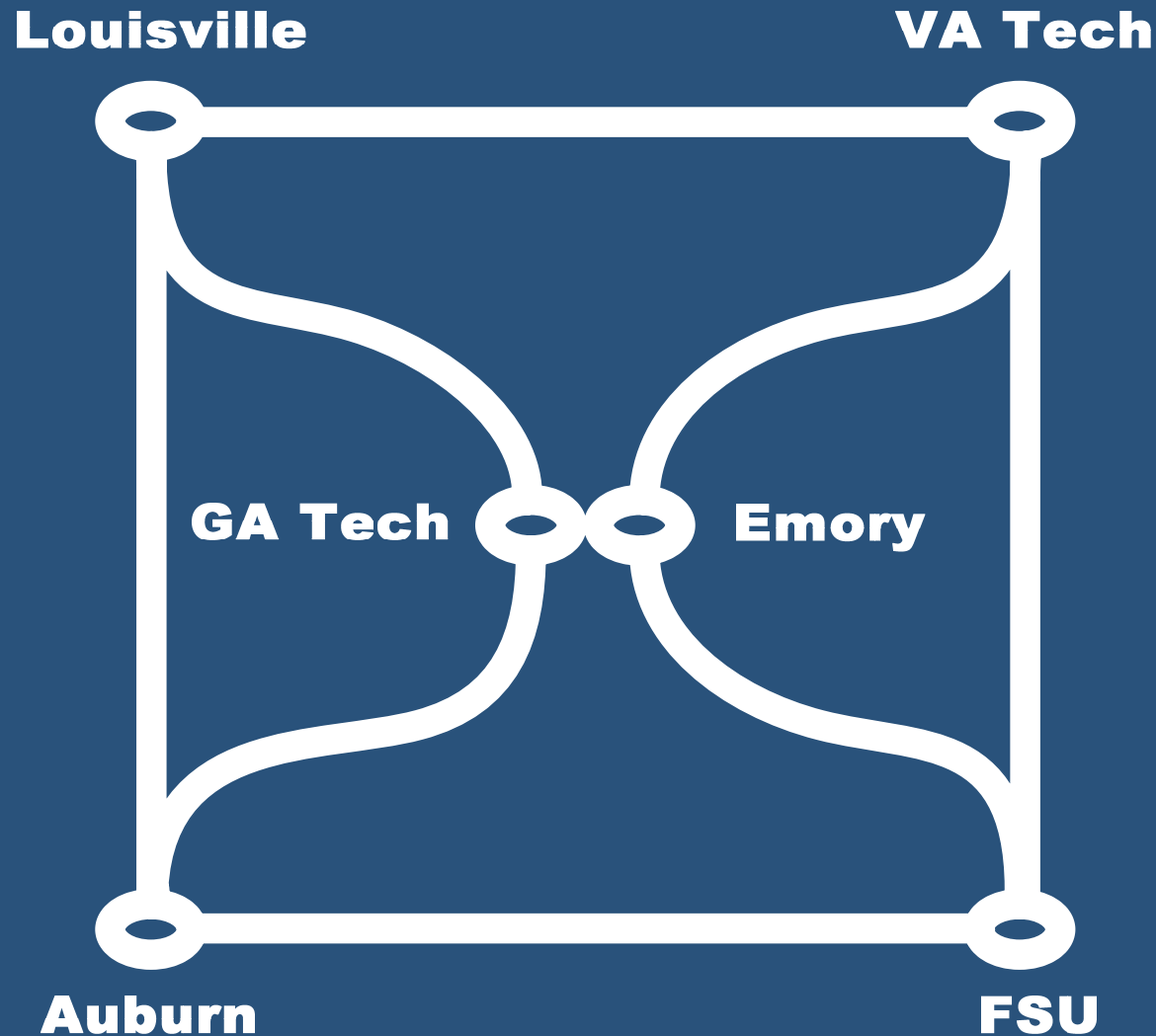
1. **Distributed Preservation Strategy**
2. **Flexible Organizational Model**
3. **Formal Content Selection Process**
4. **Capability for Migrating Archives**
5. **Dim Archiving Strategy**
6. **Low Cost to Deployment**
7. **Self-Sustaining Incentives**
8. **Simple Preservation Exchange Mechanisms with the Library of Congress**



Preservation Network Design

- Effective digital preservation succeeds by distributing copies of content in secure, distributed locations over time
- This preservation network is based on LOCKSS, a leading preservation software for distributed archiving, establishing from the beginning a distributed means of replicated archives. Our network software variant is termed **CLOCKSS** for (Collecting Lots of Copies Keeps Stuff Safe)
- Solution is low cost, open source, and securely preserves content in state that can be subsequently migrated

MetaArchive Preservation Network Partners





Dim Archiving

- Advantage: many preservation efforts conflate high accessibility online with long-term access.
- But, high accessibility equates **high costs**
- We will maintain a publicly accessible registry of collections and items preserved, but we do not intend (in this project) to provide strong public access display capabilities for the preserved content



Content Selection

- Preservation efforts are likely to be most coherent around shared focus
- Subject domain: **Southern culture and history**
- Selection of collections to be preserved made by teams of subject specialist librarians and archivists at contributing institutions
- These teams creating a conspectus of collections for consideration and prioritization
- Using collection framework of the *Encyclopedia of Southern Culture*

Conspectus Metadata Schema



- Informed by many similar efforts:
 - Dublin Core Collection Description Application Profile
 - UIUC IMLS Collection Description Metadata Schema
 - UKOLN RSLP Collection Description Schema
 - Western States Dublin Core Metadata Best Practice
- MetaArchive conspectus schema has been developed and is available on the project website:

(<http://www.MetaArchive.org>)

Cooperative Agreement Model



- Will develop a simple and flexible cooperative agreement as a model for other institutions seeking to cooperate for purposes of digital preservation
- The cooperative seeks to not only create an effective preservation network for one body of digital content, but enable the creation of many others for this important purpose



Project Timeline

Feb 2005: Conspectus complete

April 2005: Network deployment

May 2005: SW/HW tests complete

Aug 2005: Initial archiving complete

Feb 2006: Cooperative model analysis complete

Aug 2006: All subsequent project harvests complete

Jan 2007: All project goals complete



QUESTIONS ?

- Contact Info.
 - Martin Halbert - mhalber@emory.edu
- WebSite

<http://www.MetaArchive.org>