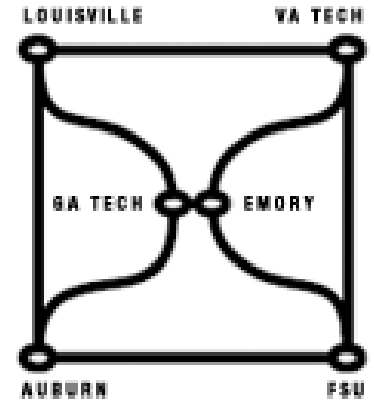


# METAARCHIVE NDIIPP Partnership



**Robert H. McDonald**

**Associate Director of Libraries for Technology & Research**

**Florida State University**

**rmcdonal@mailier.fsu.edu**

**Tyler Walters**

**Associate Director for Technology & Resource Services**

**Georgia Institute of Technology Library & Information Center**

**Tyler@gatech.edu**

**iPRES**

International Conference on Preservation of Digital Objects

*The* LIBRARY of CONGRESS

**MetaArchive NDIIPP Partnership**

**10.9.06**



# MetaArchive Project Summary

- Six partner institutions:
  - Emory - Georgia Tech - Florida State
  - Virginia Tech – Auburn - Louisville
- Collaborate w/ LC/NDIIPP – \$1.4M initial effort to develop cooperative for preservation of digital content, 2004-2010
- Origins/Precursors - MetaScholar.org:
  - AmericanSouth.org
    - Harvest metadata from cultural memory insts. pulling into central location to aggregate, index, search, and discover.  
50,000 records, +30 archives

University of Louisville

Va Tech

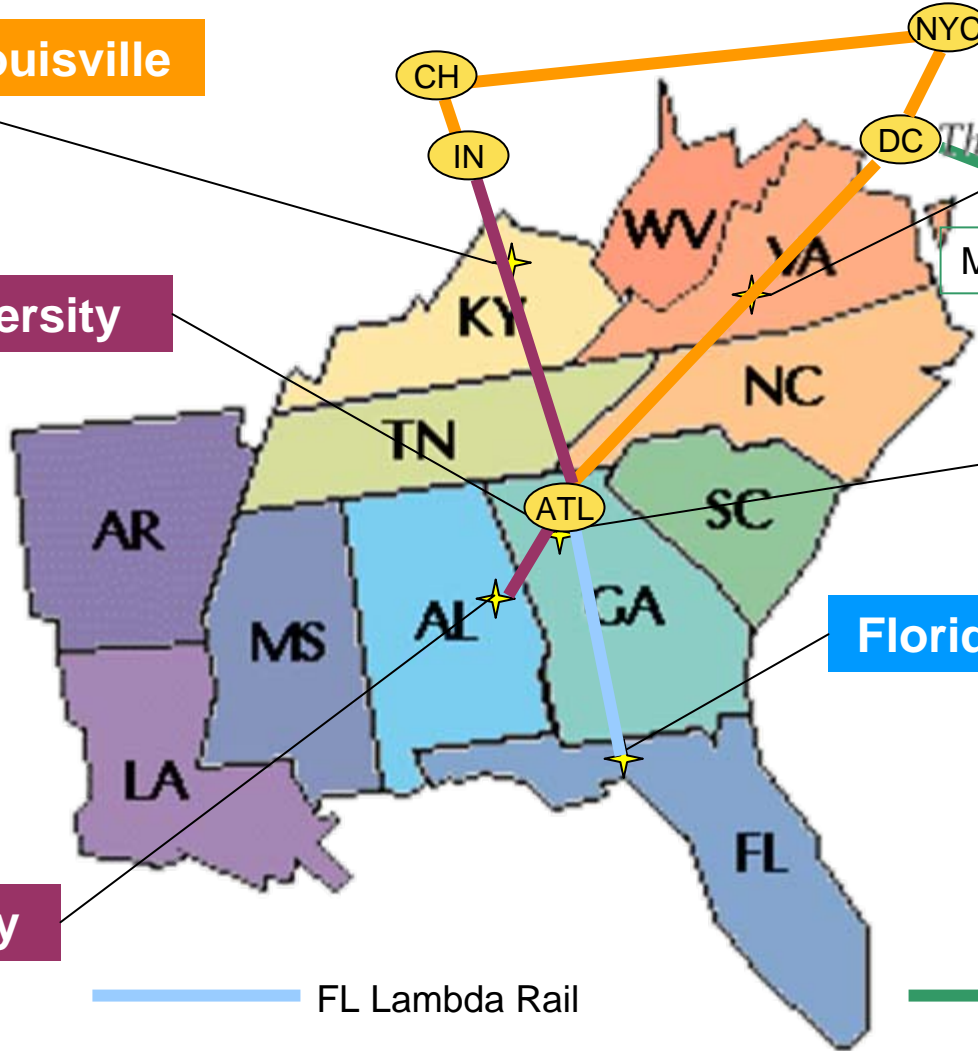
Emory University

MAX Connection to Va Tech

Ga Tech

Florida State University

Auburn University



FL Lambda Rail

MAX Network

Abilene Network

SOX Network

# iPRES

## MetaArchive NDIIPP Partnership

International Conference on Preservation of Digital Objects

10.9.06



# MetaArchive Project Goals

- 1) Create a **conspectus** of digital content held by the partner sites
  - 2) **Harvest** a body of most critical content to be preserved (3 terabytes, w/ capability to expand)
  - 3) Develop **model cooperative charter** for collaboration and sustainability
  - 4) LOCKSS-based distributed **preservation network infrastructure**  
Address geographic and institutional heterogeneity needed to safeguard each institution's digital collections
- **Distribution/Communication:** Internet2 Abilene network and local connections of the Southern Crossroads (SoX) network consortium and the Mid-Atlantic Crossroads (MAX) network consortium

# MetaArchive Deliverables to Date

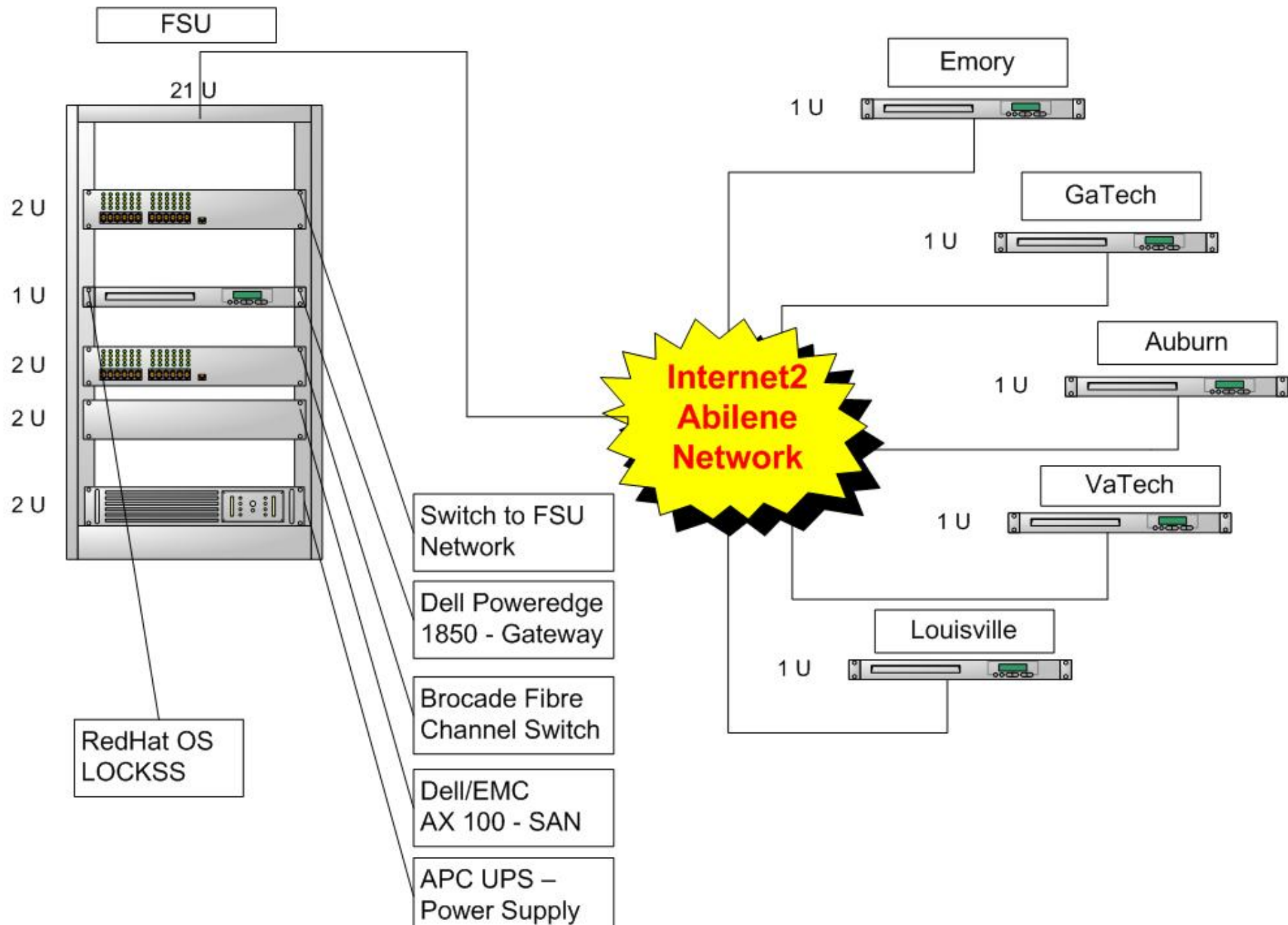
- **2004-2005**
  - Create Collection Conspectus
  - LOCKSS Customization
  - Network Setup
- **2005-2006**
  - Network Launch/Initial Harvest
  - Charter Cooperative Agreement
- **2006-2007 (projected)**
  - Non-Profit Setup (Educopia Institute)
  - Extend partnership beyond original 7 partner institutions
  - Finish Current Collection Ingest



# MetaArchive Technology

- **LOCKSS Scalability**
  - 2.2 Terabytes Storage Space Available
  - Proven Security (SE-Linux) Across Multiple Networks
  - Enterprise Installation
    - RedHat
    - CentOS
- **Database Driven Conspectus**
  - Manages Collections within LOCKSS
  - Network Administrative Management Tool
- **Collaboration Tools**
  - MediaWiki, VideoConference, Teleconference

# MetaArchive Preservation Network



# MetaArchive Conspectus Database

**Conspectus Database – Ties in with LOCKSS Central Plugin Repository**

## Log In Here

Login:

Password:

If you've lost or forgotten your password, click [Here](#).

## Login Attempt Failed

## Actions

- [View All RDF Data](#)
- [View By Institution](#)
- [View All Valid RDF Data](#)
- [View Selected Collections](#)
- [View Conspectus Schema](#)

## View Existing Descriptions

Georgia Tech ETDs [edit]  
<http://etd.gatech.edu/index.html>

**Collection ID:** 32

**Alternate Title:** Georgia Tech Electronic Thesis and Dissertation Collection

**Description:** The Electronic Thesis and Dissertation Collection at Georgia Tech covers Fall 2003 through the present. This program is jointly sponsored by the Graduate Studies and Research Office and the Library and Information Center. Georgia Tech is a member of the Networked Digital Library of Theses and Dissertations and participates in the Open Archives Initiative.

**ESC Heading:** Education

**ESC Heading:** Transportation

**ESC Heading:** Art and Architecture

**ESC Heading:** Law and Politics

**ESC Heading:** Geography

**ESC Heading:**

**ESC Heading:** Business, Agriculture, and Industry

**ESC Heading:** Social Class

**ESC Heading:** Environment

**ESC Heading:** Recreation

**ESC Heading:** History, Manners, and Myth

**LCSH Heading:** Dissertations, Academic -- Georgia --Atlanta.

**LCSH Heading:** Electronic dissertations -- Georgia --Atlanta.

**Example Conspectus Entry – GaTech ETDs**



TM

1
7
0
[Issue Tracker](#)
[Refresh Archival Units](#)

Refresh Cache:

For Caches in Network:



Filter Manager by Group:

Filter Manager by Network:


**Caches**
[View All Caches](#)
[View Down Caches](#)
[View Up Caches](#)
[View Unknown Caches](#)
[Add New Cache by Field](#)
[Add New Cache From](#)
[Email](#)

## LOCKSS ADMIN INTERFACE for METAARCHIVE NETWORK

<a href="#">Status</a>	<a href="#">Status Last Updated</a>	<a href="#">Institution Name</a>	<a href="#">IP Address</a>	<a href="#">Reverse DNS Entry</a>	<a href="#">Log Page</a>
	September 20, 2005, 12:02 pm	Louisville University		meta-vault.library.louisville.edu	<a href="http://meta-vault.library.l...">http://meta-vault.library.l...</a>
	September 20, 2005, 12:02 pm	Auburn University		meg.lib.auburn.edu	<a href="http://meg.lib.auburn.edu">http://meg.lib.auburn.edu</a>
	September 20, 2005, 12:02 pm	Emory University		ndiip.library.emory.edu	<a href="http://ndiip.library.emory...">http://ndiip.library.emory...</a>
	September 20, 2005, 12:02 pm	Florida State University		clockss.lib.fsu.edu	<a href="http://clockss.lib.fsu.edu:">http://clockss.lib.fsu.edu:</a>
	September 20, 2005, 12:02 pm	Florida State University		clockss2.lib.fsu.edu	<a href="http://clockss2.lib.fsu.edu">http://clockss2.lib.fsu.edu</a>
	September 20, 2005, 12:02 pm	Georgia Institute of Technology		ndiiplockss.library.gatech.edu	<a href="http://ndiiplockss.library.c...">http://ndiiplockss.library.c...</a>
	September 20, 2005, 12:02 pm	LOCKSS		sul-lockss27.Stanford.EDU	<a href="http://sul-lockss27.stanfo">http://sul-lockss27.stanfo</a>

metaarchive - 14:44:08 - Oct 5 02:44

**view: metaarchive**

CAMERA DESKTOP PICTURE

SIZE SHOW SETUP HELP

**robert mcdonald**  
CAM 15 f/s 313 kb/s (CIF)

298 kb/s 1 participant



RAI: metaarchive

Listen

Talk

Minimize Dock Settings Help Quit

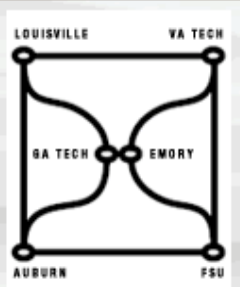
Leave this meeting

- Available
- My EVO Buddies
  - My EVO Communities
    - Internet2 (USA)
      - Meetings
      - Users
        - Ching LEE
  - Other EVO Communities
    - High Energy Physics
    - Fusion
    - National Education and Research
    - Astronomy
    - International Projects

Internet2 (USA) metaarchive EVO Message x

[14:45:55] robert mcdonald the cooperative agreement is in its 6th draft

# EVO Web Collaboration Tools



- navigation
- Main Page
  - Community portal
  - Current events
  - Recent changes
  - Random page
  - Help
  - Donations

search

- toolbox
- What links here
  - Related changes
  - Special pages
  - Printable version

- article
- discussion
- edit
- history
- move
- watch

# Main Page

- Contents** [hide]
- 1 MetaArchive of Southern Digital Culture Working Documents
    - 1.1 Conspectus
    - 1.2 Cooperative Agreement Analysis Plan (CAAP)
    - 1.3 Technology Documents
    - 1.4 NEH Grant Rough Draft
    - 1.5 IMLS Technology Training Grant Rough Draft
    - 1.6 Security and Continuity Plan Rough Draft
  - 2 Wiki Editing Help

## MetaArchive of Southern Digital Culture Working Documents [edit]

Contents of the old wiki can be found [here](#)

## Conspectus [edit]

- [ScopeDocument](#)
- [DescriptiveSchema](#)
- [DescriptiveSchemaElements](#)
- [MaterialsDocument](#)
- [SubjectsDocument](#)

## Cooperative Agreement Analysis Plan (CAAP) [edit]

**MediaWiki Portal – Document Collaboration**

# Organizational Issues

## “Sustaining the Cooperative”

### Advancing our models:

- **Organizational**
- **Technology**
- **Service**

### Organizational Model of Sustainability:

- Individual institutions: **Emory, etc.**
- Consortia-building: **Cooperative Charter**
- Non-profit mgmt. entity: **Educopia Institute**
- National strategic partnerships: **LC/NDIIPP**



# MetaArchive Digital Sustainability Model

## Economic

General Resource Inputs:

- 1) Financial
- 2) Infrastructure
- 3) Expertise

**#4 Inputs: National/International Government Agencies & Organizations**  
 More Resources, Strategic Alignment  
 High-Level Communication/Coordination  
 New Economic Models  
 International Collaboration  
 NDIPP Funding  
 DPC Projects

**#3 Inputs: Organizations, Projects, Technologies**  
 Oversee, Guide, Coordinate  
 More Strategy / More Resources  
 Sustain Innovation (Ideas/Technologies)  
 Mellon Funding

**#2 Inputs: Add & Combine**  
 Resource inputs from six universities  
 New Infrastructures  
 New Ideas / Technologies begin and broaden  
 NDIPP Funding

**#1 Base of Inputs:**  
 Universities' Infrastructure,  
 People & Expertise,  
 Initial Technology

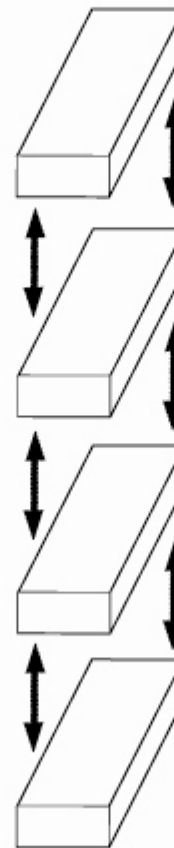
## Organization

LC / NDIIPP  
DPC

Educopia  
Institute

MetaArchive Cooperative  
(GT, FSU, Auburn, VT,  
Louisville, with Emory)

Emory University



## Technology

LC / NDIIPP  
DPC

Educopia  
Institute  
(plus, expand  
MetaArchive  
membership)

MetaArchive  
Cooperative,  
i.e. Georgia Tech,  
other developers

Stanford / LOCKSS  
Development Group

## Collections

Sustain:  
 Content (Reliability)  
 Context (Provenance)  
 Structure  
 Authenticity

Services:  
 Preservation  
(Content ingest /  
retrieval)  
 Format migration  
 Digital Collection  
 Disaster Recovery  
 Digital Preservation  
 Network Consulting  
 LOCKSS services

# MetaArchive as Consortia: Cooperative Charter

- Cooperative Charter addresses four areas:
- A) Types of partnerships in the MetaArchive Cooperative:
- **Development Partners** – responsible for developing and testing the preservation network technology and operating a preservation node
- **Preservation Partners** – responsible for operating a preservation node, ingesting collections from partner institutions, and making the node available for testing
- **Contributing Partners** – cultural memory institutions that possess digital content to preserve via the MetaArchive Preservation Network. They contribute fees for this service, do not operate a node
- **Sponsored Partners** – cultural memory institutions that own or possess digital content they wish to preserve, but cannot contribute service fees. A Development Partner sponsors the preservation of their digital content

# Cooperative Charter

## B). How MetaArchive is organized, governed, and communications

- **Steering Cmte.**
  - **Content Cmte.**
  - **Preservation Cmte.**
  - **Technical Cmte.**
- Partner institution reps can serve terms on the committees, ensuring a broad participation in governance and operations

## C). Cooperative services to members

- **Digital preservation** (network dev./maintenance, content ingest/retrieval)
- **Format migration**
- **Digital collection disaster recovery**
- **Digital preservation network consulting**
- **LOCKSS services**

## D). 1. Technical specs for preservation network

## 2. Copy of memo of understanding between the six universities

**i PRES**

**MetaArchive NDIIPP Partnership**

International Conference on Preservation of Digital Objects

10.9.06



# MetaArchive and Educopia Institute

- Non-profit management entity – Three issues:
    - 1) Continuing need for financial resources
    - 2) Expose MetaArchive to new digital projects to inform development
    - 3) Economically efficient, catalytic structure to bring these about
  - Educopia Institute:
    - Provide oversight to MA Cooperative and other digital projects
    - Low-cost, low overhead conduit for digital library, scholarly communications technology projects
    - Advance cyberinfrastructure needed to drive research, teaching and learning in contemporary digital era
    - NSF (2003) and ACLS (2006) Cyberinfrastructure reports:
      - **Scholarly activity – teaching, research, learning, knowledge transfer via scholarly communications – need rational, strategic cyberinfrastructure**
- EI: Generate DL technology projects to support mission and goals**

iPRES

MetaArchive NDIIPP Partnership

International Conference on Preservation of Digital Objects

10.9.06



# Nat'l Strategic Partnership: LC/NDIIPP

- **Access to resources, context of a nat'l digital preservation agenda**
  - LC's digital preservation partners and their approaches
  - Expertise within LC itself
  - Funding - \$650,000, 2004-2007 - servers, project staff
  - Driver to develop MA models: Organizational, Technology, Services
- **Facilitate communication on dig. sustainability, gather know-how**
- **Economic models for digital library activity**
  - Consultant, Paul Courant – Univ. of Michigan, participates in meetings and projects regarding economic sustainability issues for digital pres.
- **Align programs and projects with Nat'l/Int'l initiatives**
  - Increasingly important to generating revenues
- **Emerging Nat'l/Int'l approaches to guide funding**
  - UK: JISC and DCC
  - Australian Partnership for Sustainable Repositories [APSR]
  - Netherlands KB E-Depot
  - U.S.: NDIIPP

iPRES

MetaArchive NDIIPP Partnership

International Conference on Preservation of Digital Objects

10.9.06



# MetaArchive: Organizational Issues

## International Strategy?

- Membership in **Digital Preservation Coalition**
  - Further embed MA's activities in Nat'l/Int'l strategic initiatives

**Summary:** In two years, much work has taken place to grow the MetaArchive Cooperative into a sustainable organization:

- Four levels of organization building:
  - **Institution**
  - **Consortium**
  - **Non-profit Management Entity**
  - **National/International Strategic Partner**

...should result in a dynamic and developing organization that productively addresses distributed digital preservation issues

# Digital Preservation of Interest

- *LC NDIIPP*  
<http://www.digitalpreservation.gov/>
- *MetaArchive*  
<http://www.metaarchive.org>
- *KB e-Depot*  
<http://eprints.rclis.org/archive/00005795>
- *JISC UK LOCKSS Pilot*  
[http://www.jisc.ac.uk/index.cfm?name=programme\\_lockss](http://www.jisc.ac.uk/index.cfm?name=programme_lockss)
- *Australian Partnership for Sustainable Repositories*  
<http://www.apsr.edu.au/>
- *SDSC/NARA Chronopolis Project*  
<http://globalstor.org/pdf/presentations/Moore-chronopolis.pdf>



# Collaboration Tools of Interest

- *EVO*  
<http://evo.caltech.edu/>
- *VRVS*  
<http://www.vrvs.org/>
- *ConferenceXP*  
<http://research.microsoft.com/conferencexp/>
- *Jabber*  
<http://www.jabber.org/>
- *PHPCollab*  
<http://www.php-collab.com/>
- *MediaWiki*  
<http://www.mediawiki.org>

