

# Creating a **National Gallery** of the **Spoken Word** Online

A Digital Library Initiative  
Phase 2 Project



# Creating a **National Gallery** of the **Spoken Word** Online

Mark Lawrence Kornbluh

Presentation to the 2001 Joint International Conference  
of the Association for Computers and the Humanities and  
the Association for Literary and Linguistic Computing,  
New York University

June 14, 2001



# Creating a **National Gallery** of the **Spoken Word** Online

The National Gallery of the Spoken Word is funded by Michigan State University, it's partnering institutions, and the National Digital Library Initiative Phase II spearheaded by the National Science Foundation.



# Creating a **National Gallery** of the **Spoken Word** Online

GOAL: Develop the ability of digital libraries to preserve and deliver aural resources in a wide variety of useful formats.



# The Partners

The logo for Michigan State University, featuring the text "MICHIGAN STATE UNIVERSITY" in white, uppercase letters on a green rectangular background. The text is arranged in two lines, with "MICHIGAN STATE" on top and "UNIVERSITY" on the bottom, separated by a thin white horizontal line.

MICHIGAN STATE  
UNIVERSITY

## A. Michigan State University

1. **MATRIX:** The Center for Humane Arts, Letters, & Social Sciences
2. **MSU Libraries**
  - a. **Vincent Voice Library**
  - b. **Digital Sources Center**
3. **Speech Processing Laboratory**
4. **MSU College of Education**
5. **College of Arts and Letters**

# The Partners

## **B. Northwestern University**

1. New Media Center
2. *Oyez, Oyez, Oyez*
3. *History & Politics Out Loud*

## **C. University of Colorado, Center for Spoken Language Understanding**

## **D. The Chicago Historical Society**

# Importance of Aural Resources

1. **Speeches, interviews, and readings contain a wealth of information unavailable through text alone**
  - A. For **researchers** throughout the humanities & social sciences
  - B. For **teachers** as they bring material to life
  - C. For **students** working on class projects
  - D. For **librarians** and **archivists** as gateways to online repositories
  - E. For **general public audience** wanting a deeper understanding of important events in American politics and society



# Importance of Aural Resources

## 2. Aural resources have particular value for specific segments of the population

- A. For the **visually impaired**
- B. For **students** who are aural learners
- C. For segments of society with continuing lively oral traditions, who are often **economically disadvantaged**

# Challenges

## **Central challenges to the development of the potential of digital libraries to deliver aural resources**

1. Best Practices in Digital Sound Recording and Digitization
2. Development of a Digital Archive Structure for AV Materials
3. Selection of a Robust Metadata Format
4. Searching and Sorting Strategies/Alternatives
5. Development of Alternative Delivery Interfaces to Facilitate Diverse Users and Needs
6. Building Trans-Institutional Resources

# **Building a Robust Large-Scale Trans-Institutional Digital Archive for Aural Resources**

Stage 1: Conceptualizing a Digital Archive Infrastructure for Long-term Preservation and Flexible Access

# **OAIS: Open Archival Information Systems**

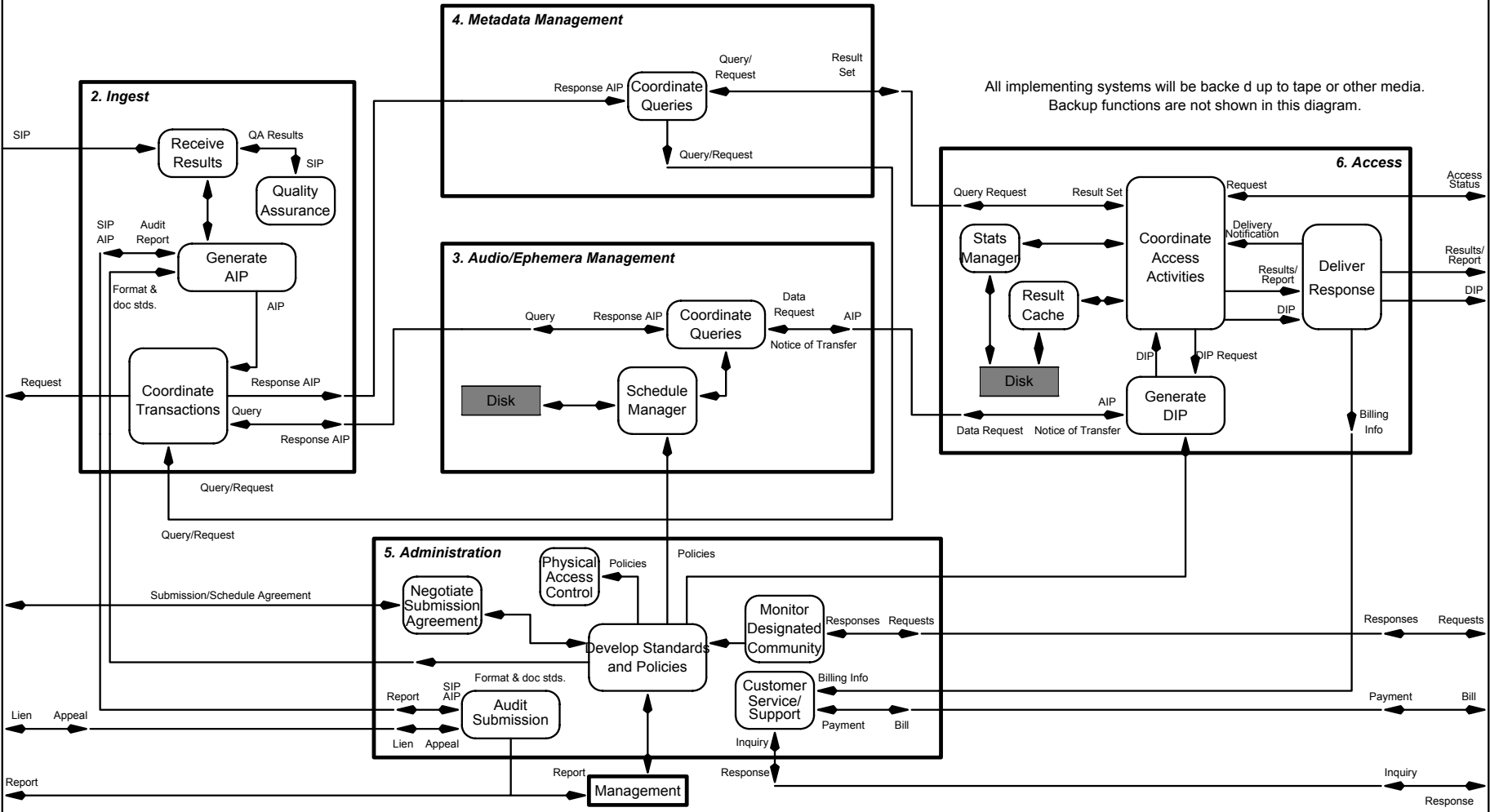
1. Methodology for creating an information ingestion, storage, and retrieval system
2. Facilitates discussion between institutions on audience, preservation techniques, and dissemination
3. An agreement between institutions that facilitates the sharing of information/files



- AIP** Information package in storage format
- DIP** Information package in distribution format
- SIP** Information package in submission format
- Designated Community** Defined group of consumers who must be able to understand the contents of the archive

P  
r  
o  
d  
u  
c  
e  
r

C  
o  
n  
s  
u  
m  
e  
r



All implementing systems will be backed up to tape or other media.  
Backup functions are not shown in this diagram.

**DRAFT**

*NGSW Delivery Archive  
Composite of Functional Entities*

Diagram by Dennis Boone and Scott Pennington, last modified 10-4-00 17:35, after OAIS Draft Recommendation diagram F-1

# **Building a Robust Large-Scale Trans-Institutional Digital Archive for Aural Resources**

Stage 2: Choose a Metadata Structure to Facilitate Long-term Preservation, Varied Access, and Interoperability

# **METS: Metadata Encoding and Transmission Standard**

1. Standard but flexible metadata schema
2. Affords the ability to share information between archives and institutions
3. Allows for the integration of legacy metadata as well as addition of new data for any object



# **Building a Robust Large-Scale Trans-Institutional Digital Archive for Aural Resources**

Stage 3: Developing Varied Delivery Interfaces for Diverse User Groups

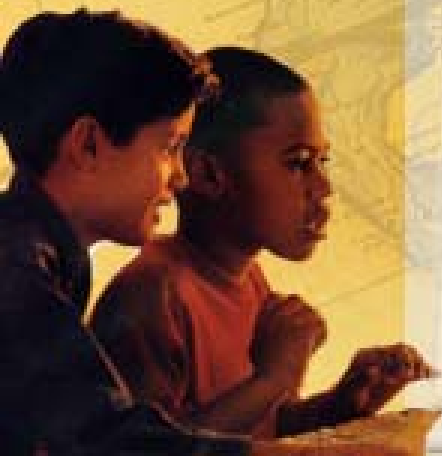
# Proposed Delivery Formats

1. Archival: Search, Browse, Sort
2. Galleries: Content-Based Groupings
3. Exhibits: Structured Presentations
4. Educational Classrooms



# Diverse Audiences for Delivery

1. K-12 classes
2. University courses
3. Scholarly research
4. Media functionality
5. Public outreach
6. Differentially-enabled users



# **Building a Robust Large-Scale Trans-Institutional Digital Archive for Aural Resources**

Stage 4: Developing Methods for User-Generated Metadata to Increase Functionality of the Archive Over Time

1. Commentary
2. Supporting Material
3. Usage Information
4. Collateral Links

# The First NGSW Galleries: Fall 2001

1. **Oldest Voices:** Selections of the Earliest Recordings from the Vincent Voice Library.
2. **The Thirties:** Social and Political Voices from the Age of the Great Depression.
3. **Most Famous Speeches:** Classic Speeches in American History.
4. **JFK, LBJ, and RMN:** Presidential Speeches and Recordings from the Sixties and Seventies.
5. **Native American Quilters:** Recorded stories of Native American quilters and Mexican American folk artists from across the Midwest
6. **The World of Studs Terkel:** Selections from Studs' Classic Radio Shows.
7. **Oral History On Line:** Best Practices and Samples for Doing Digital Oral History

# A National Gallery of the Spoken Word

1. National Project to develop best practices for digitizing storing, and delivering sound resources.
2. Development of integrated set of tools, templates, and guidelines for aural archives.
3. Creation of valuable cross-institutional online aural collections demonstrating best practices
4. Delivery of these collections in a multimedia environment that maximizes their usefulness
5. Testing & integration of collections in K-12 and university classrooms

