Yesterday’s Voices, Today’s Workflow: Oral Histories using CONTENTdm

CONTENTdm Users Group Meeting: CONTENTdm@work!
ALA Annual Conference
Anaheim, CA
June 30, 2008

Amanda A. Hurford aahurford@bsu.edu
Digital Initiatives Multimedia Developer
The Middletown Digital Oral History Project

- Library Services and Technology Act Grant funded project - $25,125
- Second consecutive LSTA digitization grant awarded to Ball State University Libraries
- Library-wide project in collaboration with Ball State University Center for Middletown Studies
Collections Digitized

- Black Muncie Oral History Collection, 1971-1978
- Black Middletown Collection, 1981
- Other Side of Middletown Collection, 2003
- Middletown Jewish Oral History Project I Collection, 1978-1979
- Middletown Jewish Oral History Project II Collection, 2002-2003
- Muncie Catholic Churches Oral History Collection, 2006
- After the grant: Muncie Labor Oral History Collection, 2005-2006
Step 1: Planning

- Across two departments
  - Archives and Special Collections:
    - Identified, processed, and inventoried collections
    - Listened to clips of interviews from each project and identified any preservation or audio quality issues
    - Located release forms and identified any potential copyright and privacy issues
  - Metadata and Digital Initiatives (MADI):
    - Researched digitization standards and equipment
    - Determined specifications for digitization
    - Got recommendations for equipment
- Funding – Wrote and secured the grant
Step 1: Planning

Archives & Special Collections
- Process collection
  - Create inventory
  - Listen to clips
  - Consider copyright/privacy

Metadata and Digital Initiatives (MADI)
- Research standards and equipment
  - Determine digitization specs
  - Recommend purchases

- Write grant/secure funding

Planning

Begin workflow
Step 2: Audio Digitization

- Used student labor to digitize over 200 interviews
- Digitized with CD quality settings
  - 16 bit depth, 44,100 kHz sampling rate
- Created several versions of each file
  - Archival WAV
  - “Cleaned” WAV (anti-hiss, volume normalization)
  - Access copy for streaming
- Enlisted help from Library Information Technology Services
Digitization Equipment

- Gateway
- Cassette Deck
- Audio Editing Software
- Firebox
- Headphones
Step 2: Audio Digitization

LITS (Techies)

- Obtain/install equipment
- Hire/train student labor

Audio digitization

Digitize

MADI

- Master WAV

“Cleaned” WAV

Access file WMA

Place on streaming server

LITS (Techies)
Step 3: Transcription

- Good: Full-text searchable
- Bad: Time and $$$
- Other options:
  - Abstracts
  - Indexes
  - Audio only
- In-house or outsource?

---

Archives and Special Collections
University Libraries

Other Side of Middletown Project
MSS154 | R. 154
Project Identification Number: Training the Young - M93TY11-1-02/21
Interviewers: Carla Burke and Carrie Kissel
Interviewee: Dr. Sam Abram
Date of interview: 2/11/2003

Abram: Where do you want me to start?
Burke: Wherever [laughter]
Abram: So the subject of our interview is what, now?
Kissel: Training the young; education as well as informal socialization of children. How children learn how to act and how not to act in important lessons is what we want to focus on in our chapter.
Abram: Well let me just start and you all interrupt at any time and you ask me questions, because I'm not—haven't prepared anything to give you or to follow us online. Let me just start by saying I think that it starts with the home and, uh, we are talking about training or educating the young; if, uh, it must start there. Some of the conditions in the home weren't other additional help in those early years.

[1:00]
Transcription Tools

- Voice Recognition Software?
- Transcription Software
  - Express Scribe
- USB Foot Pedals
- Headphones
- Format and Style Guide
  - Baylor University Institute for Oral History
    - [http://www.baylor.edu/content/services/document.php?id=1414](http://www.baylor.edu/content/services/document.php?id=1414)
Other Transcription Considerations

- Privacy
- Speech patterns and variations
- Editorial comments
- Time codes
- Format of the page
- Document decisions on a wiki

Transcription Process Management

- Split tasks
  - Transcription
  - Editing
  - Quality Control
  - Package and Finalize
- Evaluate progress and retrain if needed
Step 3: Transcription

LITS

Obtain/install equipment
Select Style Guide/Format
Train staff
Train staff
Transcribe
Transcribe
First draft
Edit/timecode
Finalize PDF

Archives/MADI
Step 4: Metadata

- Works alongside transcripts
- Provides structured description to oral histories

<table>
<thead>
<tr>
<th>Title</th>
<th>Campbell, Cordell (Audio Interview and Transcript)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interviewer</td>
<td>Vander Hill, Warren</td>
</tr>
<tr>
<td>Interviewee</td>
<td>Campbell, Cordell</td>
</tr>
<tr>
<td>Date Recorded</td>
<td>2005-12-19</td>
</tr>
<tr>
<td>See Photo of Interviewee</td>
<td><a href="http://libx.bsu.edu/MidOrHis/PHOTOGRAPHS/Labor/R76.2photo.jpg">http://libx.bsu.edu/MidOrHis/PHOTOGRAPHS/Labor/R76.2photo.jpg</a></td>
</tr>
<tr>
<td>Run Time</td>
<td>01:28:19</td>
</tr>
<tr>
<td>Location Recorded</td>
<td>Muncie, Indiana</td>
</tr>
<tr>
<td>Subject (LCSH)</td>
<td>Oral History</td>
</tr>
<tr>
<td>Industries -- Indiana -- Muncie</td>
<td></td>
</tr>
<tr>
<td>Labor unions -- Indiana -- Muncie</td>
<td></td>
</tr>
<tr>
<td>Muncie (Ind.) -- Economic conditions</td>
<td></td>
</tr>
<tr>
<td>Labor -- Indiana -- Delaware County</td>
<td></td>
</tr>
<tr>
<td>Automobile industry workers -- Indiana -- Muncie -- Interviews</td>
<td></td>
</tr>
<tr>
<td>Type (DCMI)</td>
<td>Sound</td>
</tr>
<tr>
<td>Text</td>
<td></td>
</tr>
<tr>
<td>Original Physical Format</td>
<td>Audio Cassette</td>
</tr>
<tr>
<td>Digital File Format</td>
<td>Streaming Audio (WMA)</td>
</tr>
<tr>
<td>PDF</td>
<td></td>
</tr>
<tr>
<td>Digital Identification</td>
<td>R76_002</td>
</tr>
<tr>
<td>Archival Identification</td>
<td>R76.2</td>
</tr>
<tr>
<td>Repository</td>
<td>Ball State University Libraries, Archives and Special Collections</td>
</tr>
<tr>
<td>Collection</td>
<td>Middletown Digital Oral History Project Collection</td>
</tr>
<tr>
<td>Subcollection</td>
<td>Muncie Labor Oral History Project Collection</td>
</tr>
<tr>
<td>Ordering Information</td>
<td>To order a copy, inquire about permissions, or for information about prices, contact Archives Copyright 2007, Ball State University. All Rights Reserved.</td>
</tr>
</tbody>
</table>
Planning for Metadata

- Identify sources of information
- Decide what level of information to provide
- Create a scheme using Dublin Core
  - What does the user need to know?
  - What do we need to know?
  - Identify elements
  - Create user-friendly labels
- Create a guide
  - Make rules for format, controlled vocab, etc.
Creating Metadata

- Excel spreadsheets
  - Sharing
  - Manipulation
  - Upload
- Start with inventory list created in Archives
- Create controlled vocabularies as you go (names, places, subjects, formats)
- Use transcript as full-text field
- Link to related sources
<table>
<thead>
<tr>
<th>Field name</th>
<th>DC map</th>
<th>Data type</th>
<th>Large</th>
<th>Search</th>
<th>Hide</th>
<th>Vocab</th>
<th>add field</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Title</td>
<td>Title</td>
<td>Text</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>move to</td>
</tr>
<tr>
<td>2 Interviewer</td>
<td>Creator</td>
<td>Text</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>move to</td>
</tr>
<tr>
<td>3 Interviewee</td>
<td>Creator</td>
<td>Text</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>move to</td>
</tr>
<tr>
<td>4 Date Recorded</td>
<td>Date-Created</td>
<td>Text</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>move to</td>
</tr>
<tr>
<td>5 See Photo of</td>
<td>Relation</td>
<td>Text</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>move to</td>
</tr>
<tr>
<td>Interviewee</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 Run Time</td>
<td>Format-Extent</td>
<td>Text</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>move to</td>
</tr>
<tr>
<td>7 Location Recorded</td>
<td>Description</td>
<td>Text</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>move to</td>
</tr>
<tr>
<td>8 Subject (LCSH)</td>
<td>Subject</td>
<td>Text</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>move to</td>
</tr>
<tr>
<td>9 Notes</td>
<td>Description</td>
<td>Text</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>move to</td>
</tr>
<tr>
<td>10 Type (DCMI)</td>
<td>Type</td>
<td>Text</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>move to</td>
</tr>
<tr>
<td>11 Original Physical</td>
<td>Format-Medium</td>
<td>Text</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>move to</td>
</tr>
<tr>
<td>Format</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12 Digital File</td>
<td>Format-Medium</td>
<td>Text</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>move to</td>
</tr>
<tr>
<td>Format</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13 Digitization</td>
<td>Format</td>
<td>Text</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>move to</td>
</tr>
<tr>
<td>Specifications</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14 Audio File</td>
<td>Format</td>
<td>Text</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>move to</td>
</tr>
<tr>
<td>Process History</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15 Digital Date</td>
<td>Date-Available</td>
<td>Text</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>move to</td>
</tr>
<tr>
<td>Available</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16 Digital Identification</td>
<td>Identifier</td>
<td>Text</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>move to</td>
</tr>
<tr>
<td>Identification</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17 Project Identification</td>
<td>Identifier</td>
<td>Text</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>move to</td>
</tr>
<tr>
<td>Identification</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18 Archival Identification</td>
<td>Identifier</td>
<td>Text</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>move to</td>
</tr>
<tr>
<td>Identification</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19 Repository</td>
<td>Source</td>
<td>Text</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>move to</td>
</tr>
<tr>
<td>Identification</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20 Collection</td>
<td>Relation-Is Part Of</td>
<td>Text</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>move to</td>
</tr>
<tr>
<td>Of</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21 Subcollection</td>
<td>Relation-Is Part Of</td>
<td>Text</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>move to</td>
</tr>
<tr>
<td>Of</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22 Ordering</td>
<td>Rights</td>
<td>Text</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>move to</td>
</tr>
<tr>
<td>Information</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23 Copyright</td>
<td>Rights</td>
<td>Text</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>move to</td>
</tr>
<tr>
<td>Information</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24 Full Text</td>
<td>Description</td>
<td>Full Text</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>move to</td>
</tr>
</tbody>
</table>
Step 4: Metadata

Archives

- Create initial record from inventory
- Determine/Create metadata fields and guide
- Enhance records to DC format
- Save as tab delimited file ready to upload

MADI
Step 5: Upload

- Plenty of options for audio and transcript display
  - Separate items
  - Compound object
  - Embedded players
  - Local customizations

- Custom thumbnails
Step 5: Upload

- Upload
- Enable-full text searching
- Upload PDF with tab-delimited metadata
- Embed audio file with PDF
- Assign custom thumbnail
- End workflow
Lessons Learned

- Prepare collections
- Don’t underestimate time and people needed for transcription
- Name list
- Subject headings
- Release form
- Character restrictions in transcript field
Finished Product

- [http://libx.bsu.edu/MidOrHist/midorhist.php](http://libx.bsu.edu/MidOrHist/midorhist.php)