

EVIA Digital Archive Project



Jon W. Dunn and William G. Cowan
Digital Library Program
Indiana University

IU Digital Library Brown Bag Series
November 19, 2004

EVIA Digital Archive Project



- EVIA: Ethnomusicological Video for Instruction and Analysis
- Partnership between Indiana University and University of Michigan
- Funding support from the Andrew W. Mellon Foundation

Project Background



- Originated with ethnomusicologists
- Need for infrastructure to support ethnomusicology field video
- Two goals:
 - Provide means for archiving video
 - Provide access for use in teaching/learning and research

Project Timeline



- Planning Phase 2001-2002
 - Three meetings bringing together ethnomusicologists, archivists, librarians, IT professionals, video professionals, intellectual property experts
- Development Phase 2003-2005
 - Goal: Build tools and set up infrastructure to digitize, annotate, and provide access to total of 150 hours of video from 15 contributors

Units Involved at IU



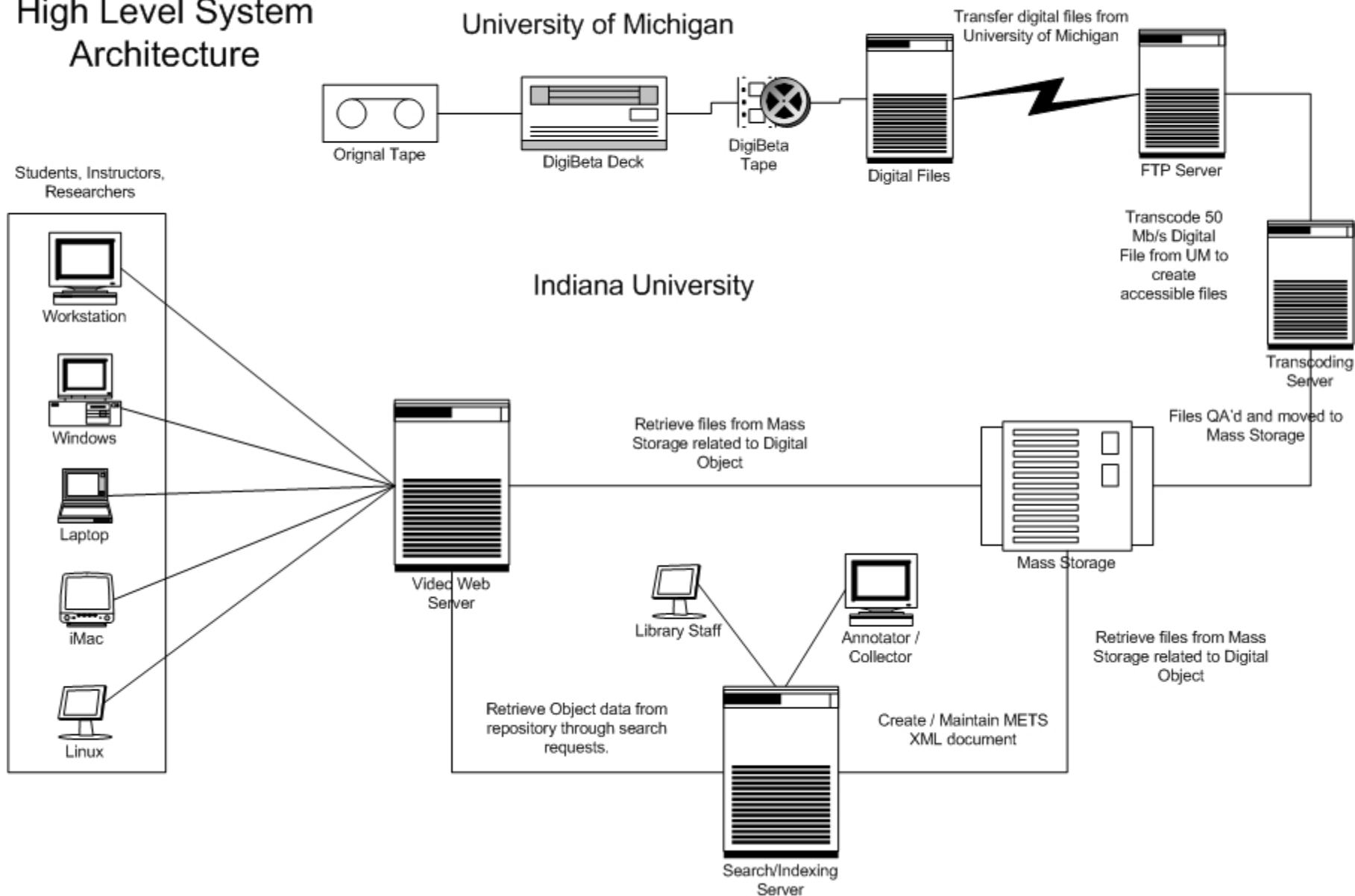
- Department of Folklore and Ethnomusicology
 - Ruth Stone
- Archives of Traditional Music
 - Daniel Reed, Alan Burdette, Suzanne Mudge, Mike Casey, Anthony Guest-Scott
- Digital Library Program
 - Jon Dunn, Will Cowan, Priya Vaidya, Nick Hansen, Jenn Riley
- UITS Digital Media Network Services
 - James McGookey

Focus of This Presentation



- Digitization, annotation process
- Tools for segmentation/annotation
- Preview of tools for access
- Lessons learned
- Future directions

EVIA Digital Archive High Level System Architecture



Application Development Environment



- Video Segmentation and Annotation application
 - Developed as Java application
 - With Borland JBuilder
 - Using an Oracle 9.2 database
 - Runs in Microsoft Windows
 - Uses QuickTime and QuickTime for Java

Metadata Standards and Schemas



- Metadata is collected by the application and stored as XML documents
 - METS 1.3 Schema
 - Structural metadata (METS structmap)
 - Container for descriptive and technical metadata
 - Pointers to digital files
 - MODS 3.0 Schema
 - Collection descriptive metadata
 - Video segment descriptive metadata

Creating the Digital Archive



- Original tape from contributor
 - Variety of formats, both analog and digital
- Dub to Digibeta tape
- Digibeta tape to 50 Mbps MPEG-2 digital file
- 50 Mbps digital file transcoded to MPEG-4 digital file
- 50 Mbps digital file stored in IU's Massive Data Storage System
- Digibeta tape stored at IU Archives of Traditional Music (ATM)
- If collector chooses, original tape may be stored at ATM as well

Creating the Digital Archive: Some Statistics



- Size of 50 Mbps digital master file
 - 10 - 15 gigs per hour of video
 - 100 – 200 gigs per collection
- Effort to create transcoded file
 - 8 hours per hour of video
 - Includes: dubbing digibeta tape, MPEG-2 encoding, checksum generation, transmission and transcoding

EVIA Digital Archive Summer Institute 2004



- June 9 – 19, 2004 at Indiana University
- Collaboration of ethnomusicologists, technicians, programmers, catalogers
- Contributions of 10 hours of video from 11 ethnomusicologists spanning 4 decades of field work
- Ethnomusicologists segmented and annotated digital video, participated in workshops, presented their results

EVIA Digital Archive Controlled Vocabulary



- Predefined categories chosen by ethnomusicologists and catalogers
- Predefined values based on authority sources chosen by catalogers
- Predefined values stored in tables in Oracle database and available to be selected in annotation application
- As needed, new values would be added by catalogers during the Summer Institute

EVIA Digital Archive Controlled Vocabulary



- Language
- Culture Group
- Geographic Location
- Genre
- Instruments
- Venue
- Theme
- Performance Type
- Participant and Participant Role

EVIA Digital Archive

Controlled Vocabulary Authorities



- Library of Congress Subject Headings
- MARC Code Lists for Relators
- Getty Art and Architecture Thesaurus
- Getty Thesaurus of Geographic Names

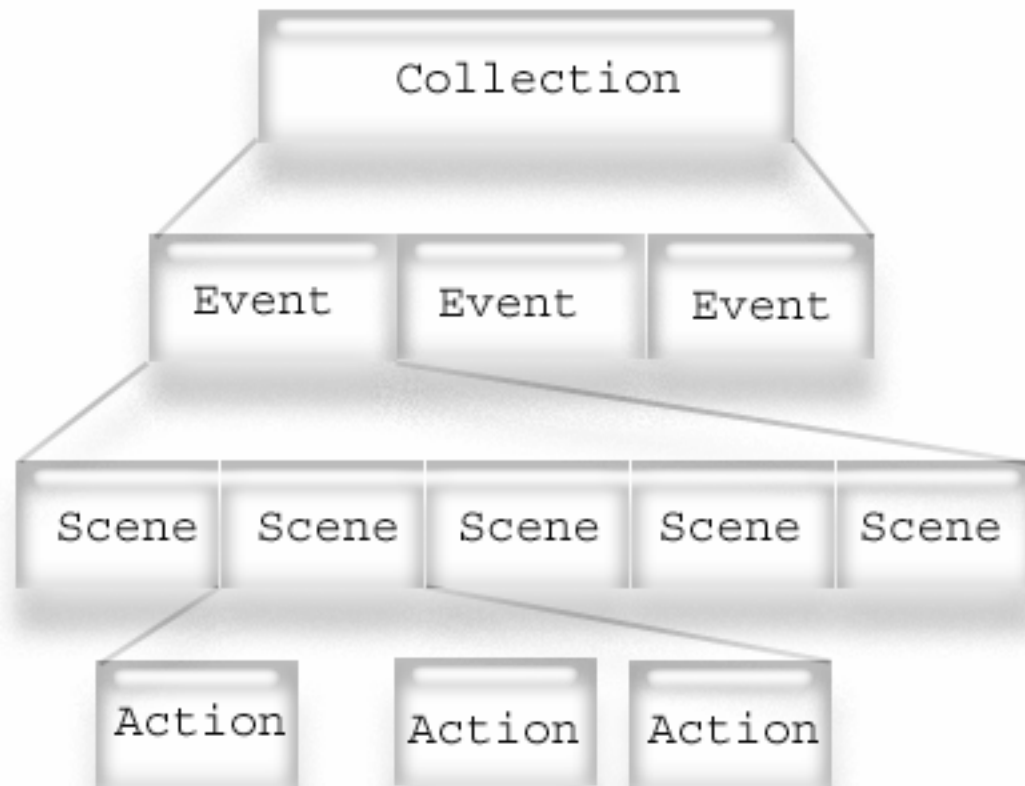


Language	Participant Role	Performance Type	Theme	Venue Type & Venue
Authority	Culture Group	Genre	Geography	Instrument



	Authority ID	Authority Desc
1	37	//delhigovt.nic.in/newdelhi/dept/obc/CC.doc
2	40	//tribal.nic.in/ragasthanareas.html
3	25	Getty Art and Architecture Thesaurus
4	26	Getty Thesaurus of Geographic Names
5	27	Hornbostel & Sachs
6	29	Library of Congress Name Authority File
7	24	Library of Congress Subject Headings
8	28	Local
9	31	MARC Code Lists for Relators, Source, Description Conventions
10	30	Pending
11	33	Pending (Garland)
12	32	Pending (Modern Language Association Thesaurus)

Video Segmentation Hierarchy



Creating Video Segments and Annotations



- Creation of initial METS XML file for each ethnomusicologist
 - Basic Data
 - Points to Transcoded Files on file server
- Event segmentation
 - Continuous video stream of all digitized video for the contributor
 - Contributor Segments continuous video stream into significant events.

Collection Workspace

Events Defined in this Collection

Simba Theatre Performing at Nyumba ya Sanaa
 Tanzanian One Theatre (TOT) Christmas Day Performance
 Ndala Kasheba and Kassongo Mapili Clayton in an Interview and Performance
 Mr. II Album Release Concert
 Hamza Kalala Interview and Performance
 St. Joseph's Choir
 African Stars Band Performs in Mhazi Mmoja
 Mvita Dancing Troupe Perform Mchiriku Music
 Classic Dansi Music Performed by Ndala Kasheba, Nguzu Viking, and King Kiki
 Rappers, R&B Artists, and Ngoma Groups Perform in a Nightclub

Segmentation Options

Create Events

Create Scenes

Collector Information

Name: Perullo, Alex

Short Biography: is usually sung in Swahili by several singers along with musicians performing on electric guitars, drums, electric bass, and keyboards.
 5. Derived from a traditional ngoma, mchiriku preceded hip hop as a popular music among youth. Often consisting of traditional drums and a keyboard, the music flows in

Short Collection Description

Alex Perullo is an assistant professor at Bryant College. He received his Ph.D. from the Department of Folklore and Ethnomusicology at Indiana University in 2003. The video deposited in the EVIA Digital Archive is drawn from his dissertation research in Dar es Salaam, Tanzania, carried out between 1998 and 2002. Many details about the history and current situation of Dar es Salaam popular music scene can be found in his dissertation "The Life That I Live: Popular Music, Agency, and Urban Society in Dar es Salaam, Tanzania." Perullo is currently working on an ethnography of Tanzanian popular music, and has written several articles on Tanzanian hip hop and copyright issues.

Collection Level Controlled Vocabulary

Geography	Genre	Instrument	Participant	Theme
Summary	Language	Performance	Venue	Culture

Refresh Summary

GEOGRAPHY:
 Tanzania, Dar es Salaam Region, null, Dar es Salaam,

INSTRUMENTS:
 Electrophones, Programmed (Electrophone), Boom box, Redi
 Aerophones, Blown (Aerophone), Organ, Kinanda,
 Electrophones, Programmed (Electrophone), Drum machine,
 Aerophones, Blown (Aerophone), Trumpet, Tarumbeta, trump
 Chordophones, Struck / Plucked (Chordophone), Bass guitar,
 Chordophones, Struck / Plucked (Chordophone), Electric gita
 Membranophones, Struck (Membranophone), Ngoma (Drum)
 Membranophones, Struck (Membranophone), Drum set, Drum

Events in this Collection

Event 1	Simba Theatre Performing at Nyumba ya Sa... [00:00:00:000 - 01:02:29:993]	01:02:29:993
Event 2	Tanzanian One Theatre (TOT) Christmas D... [01:02:29:993 - 02:46:41:940]	01:44:11:947
Event 3	Ndala Kasheba and Kassongo Mapili Clayt... [02:46:41:940 - 03:34:04:502]	00:47:22:562
Event 4	Mr. II Album Release Concert [03:34:04:502 - 04:33:39:625]	00:59:35:123

Descriptive Metadata

Heading

Simba Theatre Performing at Nyumba ya Sanaa

Description

Simba Theatre is a popular group that performs ngoma music from all over Tanzania. In this video, they perform for an audience of mostly foreign tourists.

Details

Controlled Vocabulary

Geography	Genre	Instrument	Participant	Theme
Summary	Language	Performance	Venue	Culture

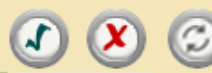
Refresh Summary

Membranophones, Struck (Membranophone), Ngoma (Drum), Ngon
 Idiophones, Struck (Idiophone), Rattle (Musical instrument), Kayamb
 Idiophones, Struck (Idiophone), Xylophone, Marimba (chondo),

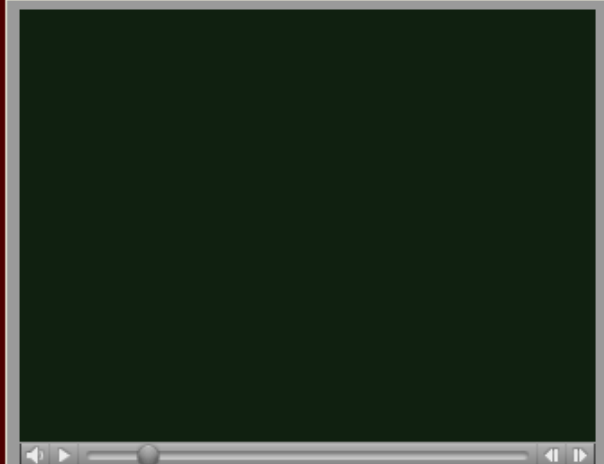
Event Time Span

Event 1

[00:00:00:000 - 01:02:29:993]



Change

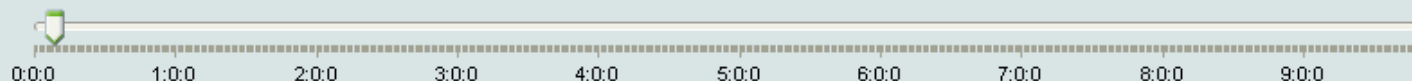


Now Playing: Event 1

00:08:06:781



Collection Timeline



Creating Video Segments and Annotations



- Contributor now segments events into scenes and actions
 - Open an Event
 - Segment as Scene
 - Descriptive text
 - Controlled vocabulary
- Segment Scene into Actions

Scenes and Actions in this Event

Scene 1	Practicing Before Concert	[00:00:00:000 - 00:03:23:183]	00:03:23:183
Scene 2	Introductory Song with Drums and a Marimb...	[00:03:23:183 - 00:07:40:970]	00:04:17:787
Scene 3	Performance of Mask and Stilt Dancing	[00:07:40:970 - 00:13:02:930]	00:05:21:960
Scene 4	Courting Dance	[00:13:02:930 - 00:16:37:872]	00:03:34:942
Scene 5	Group Ngoma Dancing	[00:16:37:872 - 00:24:23:009]	00:07:45:137

Descriptive Metadata

Heading Practicing Before Concert

Description The musicians of Simba Theatre warm up on their instruments before the concert begins. |

Details

Controlled Vocabulary

Geography	Genre	Instrument	Participant	Theme
Summary	Language	Performance	Venue	Culture

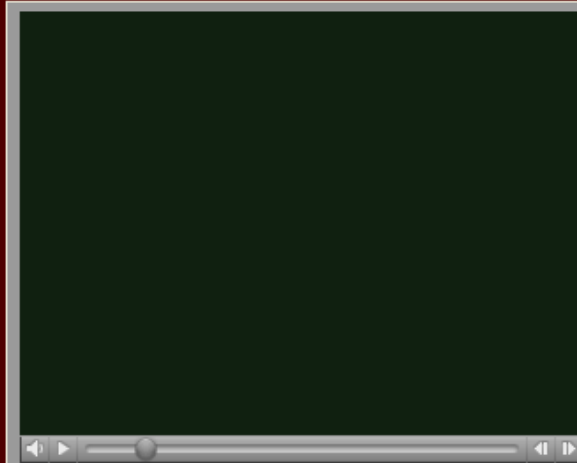
Refresh Summary

INSTRUMENTS:
 Membranophones, Struck (Membranophone), Ngoma (Drum), Ngon
 Idiophones, Struck (Idiophone), Rattle (Musical instrument), Kayamb
 Idiophones, Struck (Idiophone), Xylophone, Marimba (chondo),

Scene or Action Time Span

Scene 1

[00:00:00:000 - 00:03:23:183] Change



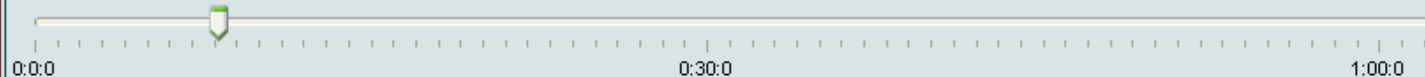
Now Playing: Scene 3, Action -

00:08:10:427

A S

Play, Stop, Previous, Next, Full Screen, Volume

Event Timeline



EVIA Digital Archive

Demo



- [EVIA Digital Archive Demo](#)
- [Sample METS and MODS XML](#)
 - METS Header / MODS Collection
 - METS File Section
 - METS Struct Map
 - MODS Descriptive Metadata
 - MODS Controlled Vocabulary

EVIA Digital Archive

Current Development



- Web access to annotated video segments
 - [EVIA Digital Archive Web Access Prototype](#)
- Storing and Retrieving XML Documents from Oracle 9.2 XDB (native XML database)
- Indexing, searching and retrieving XML from the XDB
- Interaction between Archives of Traditional Music MARC cataloging and EVIADA MODS metadata
- Tools for technical/administrative metadata capture

Lessons (Being) Learned



- A lot of effort to create digital video
 - Highly skilled work
 - Need for clearly defined production processes and procedures
- Goals of preservation and access sometimes in conflict
 - Collection development policies
 - Metadata
 - Digitization

Lessons (Being) Learned 2



- True digital video preservation not yet possible
 - No established standards or good format choices
 - Compromises on video formats necessary
 - See Jerry McDonough's presentation from DLF Fall Forum 2003:
 - <http://www.diglib.org/forums/fall2003/fallforum03.htm#p13>
- Much more work needed on technical/administrative metadata
 - Technical attributes of original source and physical and digital derivatives to support management and future migration
- Collaboration with other audio and video DL projects at IU and elsewhere

Lessons (Being) Learned 3



- More discussion
 - Evolving working relationships between scholars, librarians, archivists, technologists, video engineers

Possible Future Directions



- Tools and Technology
 - Open Source METS/MODS Video Segmentation/Annotation Tool
 - Field version of Segmentation/Annotation Tool
 - Further work with native XML databases
 - Integration into teaching/learning environments, course management systems
- Sustainability
 - Examining options: IU ATM, institutionally distributed, new nonprofit organization?
- Proposal submitted to Mellon for Interim Development Phase through June 2006.

For More Information



- Jon Dunn: jwd@indiana.edu
- Will Cowan: wgcowan@indiana.edu
- <http://www.indiana.edu/~eviada/>