AILLA: The Archive of the Indigenous Languages of Latin America

Heidi Johnson
The University of Texas at Austin
Latin American Digital Library Initiative, March 24, 2011
Acknowledgements

AILLA is a joint effort of:

- Anthropology: Joel Sherzer
- Linguistics: Anthony C. Woodbury
- Digital Library Services: Mark McFarland

AILLA is supported by:

- The College of Liberal Arts
- The University Libraries
- The National Endowment for the Humanities
AILLA's origins

- Conferences at UT in the '80's & '90's
- Discuss need to preserve Native American literatures and make them accessible.
- Sherzer & students built a pilot with seed money from COLA.
- Technical support and guidance from DLSD.
Building the permanent site

Grants from NSF & NEH for $468,973.

Hire professional staff:
- Heidi Johnson, general manager
- Erik Grostic, programmer

Hire graduate students:
- Will Fairchild, media tech extraordinaire
- Lynda DeJong (now Boudreault), archivista
- Ajb'ee Jiménez, website translator
www.ailla.utexas.org

Parallel interfaces in English and Spanish.
Online administrator metadata interface.
Online depositor metadata interface.
Downloadable metadata forms & templates.
Since 2001, AILLA has won ~$1.4 million in grant funds from NEH & NSF.
AILLA's Mission

 Preservation: irreplaceable recordings in endangered languages are permanently housed on library servers and tapes.

 Access: flexible protocols support easy access to public materials and secure restrictions for sensitive materials.
Archive statistics 1

- 230 languages from 24 countries (+15)
- 8,378 resources (+1461)
- 98 depositors (+9)
- 3202 registered users (+397)
- 69,385 media files
- 3.6 T of archival objects (+1.4T)
Archive statistics 2

Audio:
- files: 11812 (+ 3272)
- length: 4589:09:30 (+ 1037 hrs)

Video:
- files: 1106 (+ 705)
- length: 512:38:54 (+ 303 hrs)
Archive statistics 3

Text:
- pages, digital: 50,692 (+6834)
- pages, manuscript: 31,360 (+2330)

Images: 2904 (+1986)

26% restricted access files (unchanged)
Preservation

- Archive housed on library computers, maintained by UT digital librarians.
- Presentation objects accessible online; archival objects stored on tape with an (evolving, mediated) access protocol.
- Analog media are permanently housed at the Benson Latin American Collection.
What we preserve

- Multi-media resources:
  - Audio & video recordings
  - Texts: digital and manuscript
  - Images: photos, drawings, maps

- Types of resources
  - Recordings of discourse in any genre
  - Transcriptions, translations, annotations, etc.
  - Literature, oral or written
  - Analyses, lexicons, notes, sketches, etc.
  - Out of print & hard to find publications
Formats

Archival formats:
- audio: wav, 44.1 Khz, 16 or 24 bits
- video: mpeg-2
- archivable digital text: xml, txt, html
- non-archivable digital text => pdf/a
- manuscript text & images: tiff or jpg

Presentation formats:
- audio: mp3
- video: mp4
- text: pdf
- images: jpg
Metadata

Resource = a bundle of files related in terms of intellectual content, e.g. recording + annotations

Information about:
- Depositor: contact info
- Project overview, sponsor, bibliographies.
- Participants: role, demographic data
- Media: provenance, formats, relations
- Content: title, genre, description, date, place created
- References: publications
Infrastructure

MySQL database
- 44 tables, ~ 3 MB

PHP interface
- 3 sides: user, admin, depositor
- user side: login, browse/search, access
- admin: data entry, stats, error checking
- depositor: access only their materials
Security: levels of access

Provides speakers and depositors finely-grained control over resources.

Four levels:

1: public access
2: password-protection, with an optional hint
3: time-limit: specify date when file goes public
4: contact depositor or other designated controller
Security II

- Access levels assigned to individual files.
- Depositors can change settings at any time.
- Depositor control allows owner to know who is using their resources.
- Passwords allow fine or coarse control.
- Time limits can protect materials for eg a speaker's lifetime.
www.ailla.utexas.org

Comments gladly received at
ailla@ailla.utexas.org