Endangered Languages and Web-Based Archiving

Megan J. Crowhurst
The University of Texas at Austin & CELP
Contributors: Chris Beier, Heidi Johnson, Lev Michael
LSA Symposium:
The Open Language Archives Community
4 January 2002

Today's comments

- Emphasize the importance of archiving endangered language data.
- Discuss advantages of archiving EL data for various interested groups.

Advantages of data archiving

- For the field worker
- For communities providing data
- For funding agencies
- For the scholarly community

Advantages of archiving for field workers

- You know where the data is.
- Data is accessible from any web-linked site.
- You can control access to the data by others.

Advantages of archiving for field workers

Archiving is easy.
- Digital files can be uploaded to archives by FTP.
- AILLA accepts data in any format.
- AILLA staff will digitize and upload data for users.

Advantages of archiving for field workers

- Archiving is a way of publishing data.
- Prestige problem:
  - We aren't doing well as a field in recognizing the value of archiving as a publication-related activity.
    - It's new
    - No peer review process - quality control issues

Advantages of archiving for field workers

- No limit on storage capacity.
- It will be possible to
  - archive documents at various stages of processing.
  - work online using powerful browse and edit tools.

Summary:
Advantages for field workers

- Secure, accessible location
- Unlimited storage potential
- Controlled access
- Archive staff will digitize
- Online tools for analysis

Advantages of archiving for EL communities

Easy and cheap access to data.
- Access costs the community only the price of internet access.
- Archives are maintained by institutions, costing the community nothing.

Advantages for EL communities

Native speaker input.
- Checks on quality/accuracy.
- Speakers can add substantially to the content of the archive (e.g. stories).
- Speakers can contribute to analyses.

Advantages for EL communities

Long term audio storage
- EL communities can dub their own tapes from downloaded sound files.
- Particularly important in communities with low literacy.

Advantages for EL communities

Long-term storage for written materials
- Unpublished data by scholars who have not returned to the community.
- Out-of-print materials.
- Materials that are not accessible for economic or geographical reasons.

Advantages for EL communities

- For example:
  - Dictionaries
  - Grammars
  - Poetry
  - Collections of stories
  - Schoolbooks

Advantages for EL communities

The archive can be used as a base for publishing.

- Written materials can be printed in small quantities as needed.
- Useful for school materials, that need to be replenished on a regular basis.
- Materials production is placed in the hands of the EL community.

Advantages for EL communities

Archiving makes it possible for contributing communities to share resources.

- Models for grammars and dictionaries can be compared and evaluated.
- Archived lesson plans can be adapted by members of other EL communities.
  - Then the new, adapted version could be archived.

Advantages for EL communities

Confers prestige on the language.

Summary:
Advantages for EL communities

- Long-term, safe, economical storage.
- Opportunities to contribute directly to the database.
- The archive as a publishing base.
- Hub for joint projects and shared resources.
- Enabling.
- Confers prestige.

Advantages of archiving for funders

- Public record of funded research.
- Preservation and dissemination of data collected as part of funded research.
- Can help to avoid duplication of efforts.

Advantages for the scholarly community

- Data is available to anyone for sanctioned uses.
- This can make possible new kinds of archival research.

Advantages for the scholarly community

Easy access to AUDIO data
- makes possible the analysis or reanalysis of old data.
- Permits students and other scholars who can't travel to EL communities to contribute to EL work.

Advantages for the scholarly community

- Databases allow for "distributed scholarship".
  - Scholars can share access to common datasets and can collaboratively produce analyses.
  - Collaborative efforts should also be possible between scholars and members of EL communities.

And finally…

…a benefit for everybody
Archived data
is SAFE data

Archived data is protected

- from natural deterioration
- from catastrophic loss
  - accidental
  - political

Thank you.