Outreach

Jeff Good
UC Berkeley

OLAC's Needs

- Maximal involvement from the whole community
  - The more data providers involved the more useful the services become
  - More data providers mean more input on how to improve standards and services

The average user's needs

- Documents making minimal use of technical vocabulary
- Services making OLAC participation straightforward
- A central location to access all OLAC services

Impediments from OLAC's perspective

- Old habits
- Lack of awareness of the basic issues
- Idea that it is "someone else's problem"

Impediments from user's perspective

- Technical issues not directly related to linguistics (or other fields of study)
- Time and work involved

The current state of outreach

OLAC's Needs

- Maximal involvement from the whole community
- Archives have largely been self-selecting, either by direct involvement with OLAC or by already making some sort of metadata available

The average user's needs

- Documents making minimal use of technical vocabulary
- "A gentle introduction to metadata"

Aside: A gentle introduction to XML (from the TEI)

XML is an extensible markup language used for the description of marked-up electronic text. More exactly, XML is a *metalanguage*, that is, a means of formally describing a language, in this case, a *markup* language. Historically, the word *markup* has been used to describe annotation or other marks within a text intended to instruct a compositor or typist how a particular passage should be printed or laid out. Examples include wavy underlining to indicate boldface, special symbols for passages to be omitted or printed in a particular font and so forth. As the formatting and printing of texts was automated, the term was extended to cover all sorts of special codes inserted into electronic texts to govern formatting, printing, or other processing.

The average user's needs

- Services making OLAC participipation straightforward
- Vida, ORE, Viser
- EMELD

The average user's needs

- A central location to access all OLAC services
- OLAC? Linguist List?

Assessment of what is needed

Assessment of what is needed

- Contacting archives
  - Individually by members of OLAC
  - More broadly via public forums (LSA, Linguist List, . . .)

Assessment of what is needed

- Documents
  - Making the production of a non-technical documents part of the OLAC process
  - Published documents in journals giving an overview of OLAC for different communities (Language, IJAL...)
  - A FAQ on data archiving, annotation, and access in the digital age (with lots of examples)

Assessment of what is needed

- Services
  - Input from as wide a range of potential users as possible
  - Services not made public until "useful"
  - A group of official consultants from important linguistics subcommunities, preferably people not already closely involved with OLAC

Assessment of what is needed

- Central location
  - OLAC site for technical reference?
  - Linguist List for general community reference?

Overcoming the impediments

Impediments: OLAC's

- Old habits
- ??
- Lack of awareness of the basic issues
- Documents (online and in print)
- Idea that it is "someone else's problem"
- Endorsement (LSA and other high-profile organizations) (cf. 1992 *Language* article on endangered languages)

Impediments: User's

- Technical issues not directly related to linguistics (or other fields of study)
- Clear separation of general audience documents from technical documents
- Time and work involved
- Better tools, services
- Change of culture

Proposal:
Outreach Working Group

Outreach working group

- Contact archives
- Oversee creation of non-technical documents
- Work towards gaining endorsements