OLAC
Open Language Archives Community
Martin Wynne, Oxford Text Archive, Oxford University

OAI Workshop, Geneva, 12th February 2004
With a lot of help from my friends…

Steven Bird
Gary Simons
Chu-Ren Huang

Melbourne
SIL
Academia Sinica
Open Language Archives Community

Advisory Board: 15 members
Coordinators: Steven Bird & Gary Simons
Council: 7 members
Participating Archives

Over 25 Archives and Services:

http://www.language-archives.org/OLAC/repositories.html
Oxford Text Archive

- Part of Arts and Humanities Data Service, specifically *AHDS Literature, Languages and Linguistics*
- Funded as a national service by JISC and AHRB
- An archive of literary and linguistic digital resources
- Sharing resource descriptions with:
  - Other AHDS centres
  - Oxford libraries and portals
  - Subject-specific meta-archives, e.g. OLAC
Common Resource Descriptions

- OAI uses Dublin Core metadata:
- OLAC adds elements specific to community:
  - `olac:linguistic-type`:
    - lexicon, primary_text, language_description
  - `olac:language`
  - And defines controlled vocabularies
Solving the Language Identification Problem

- olac:language
  Provides codes for identifying all known languages, both living and extinct, includes three sets of unique codes
  - Unambiguous ISO 639-1 Codes  ex. en
  - Unambiguous ISO 639-2 Codes  ex. tur
  - Ethnologue Codes
    ex. x-sil-TRK

Note: ISO 639 is a subset of Ethnologue codes
Data provider approach 1: Implement CGI interface
Data provider approach 2: Export to XML repository

Usage Infrastructure

- User
  - Search Form
  - Search Engine
  - Harvested Metadata
  - Controlled Vocabulary Pages
  - Controlled Vocabulary Server

Contributor
- Metadata Entry Forms
- Repository Editor
- Shared Metadata Repository

On User Machines

Technical Infrastructure

- Metadata Harvester
- Data Provider Interface
- Archive1 Metadata Repository
- Virtual Data Provider
- Archive2 Repository in XML
- Archive2 Repository

On Central Machines

- On Archive Machines

Martin Wynne: OLAC
Data provider approach 3: Use a forms-based editor
Search all OLAC repositories:
www.linguistlist.org/olac/
Controlled vocabulary servers: e.g. www.ethnologue.com

Martin Wynne: OLAC
For more…

- Use the combined catalogue: http://linguistlist.org/olac/
- The OLAC-General mailing list: http://www.language-archives.org/
- Become a data provider: http://www.language-archives.org/docs/implement.html