Archive access and accessibility: a progress report on social networking at work

David Nathan
Endangered Languages Archive
School of Oriental and African Studies
University of London

CRASSH, University of Cambridge Dec 10, 2010





School of Oriental and African Studies, U London

Hans Rausing Endangered Languages Project

funds

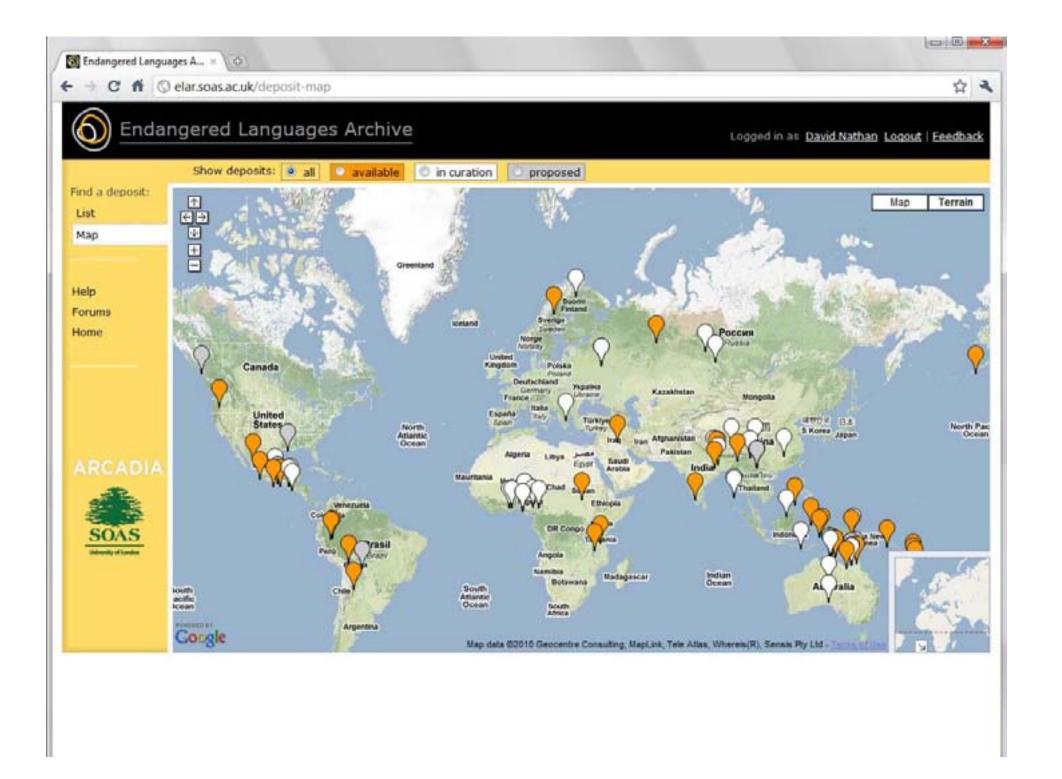
ARCADIA

Documentation Program ELDP

Archive Program

ELAR

Academic Program





The heart of providing and managing access?

- online
- free of cost
- standardised metadata
- single search portal for federated archives
- till now, access has largely been seen as
 - online resource discovery via standardised metadata
 - simple identity-driven access control



Language documentation

creation, annotation, preservation, and dissemination of a representative and multipurpose record of a language with emphasis on natural language usage in social contexts

cf Woodbury 2010, Himmelmann 2006

Stakeholders in language documentation

documenters documentation projects language consultants What is here? institutions communities heritage research curiosity journos applied users



Characteristics of language documentation

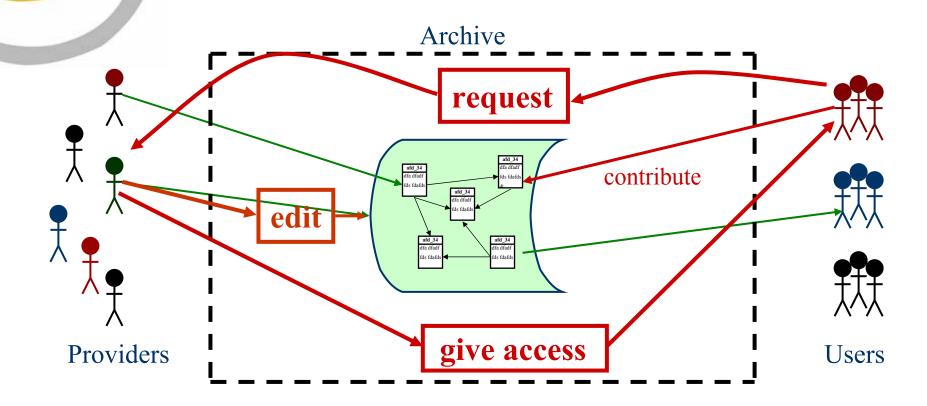
- diversity projects have different goals, methods, contexts, languages, cultures, communities, individuals
- protocol recordings and analysis of language usage in often sensitive contexts
 - more people, more genres, more private conversations, less researcher knowledge



What is a digital EL documentation archive?

- more than a data repository
- fulfilling key needs of endangered languages stakeholders:
 - diversity
 - protocol
- a forum/platform for building and conducting of relationships between data providers and users
- level the playing field between researchers and community members

ELAR - architecture





ELAR's access system

- using URCS system
- system matches a resource's URCS value with the URCS rights of a given user
- interface always shows resource and user protocol
- URCS applies at deposit, bundle or file level
- depositors can manage access control and levels

U R C S - all registered users can access.

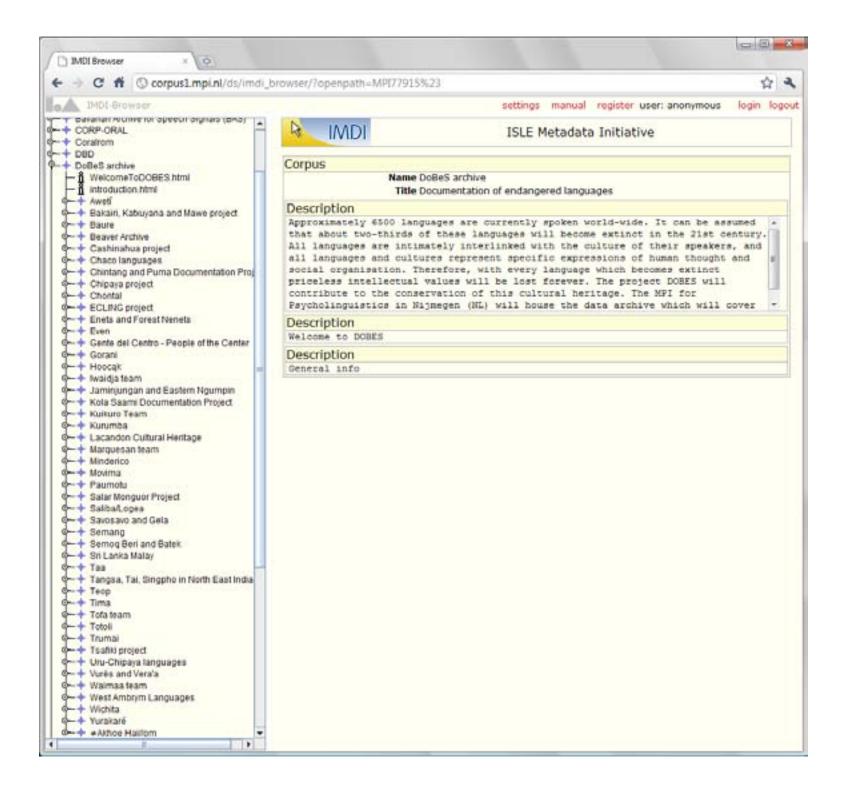
U R C S - researchers and community members are allowed access

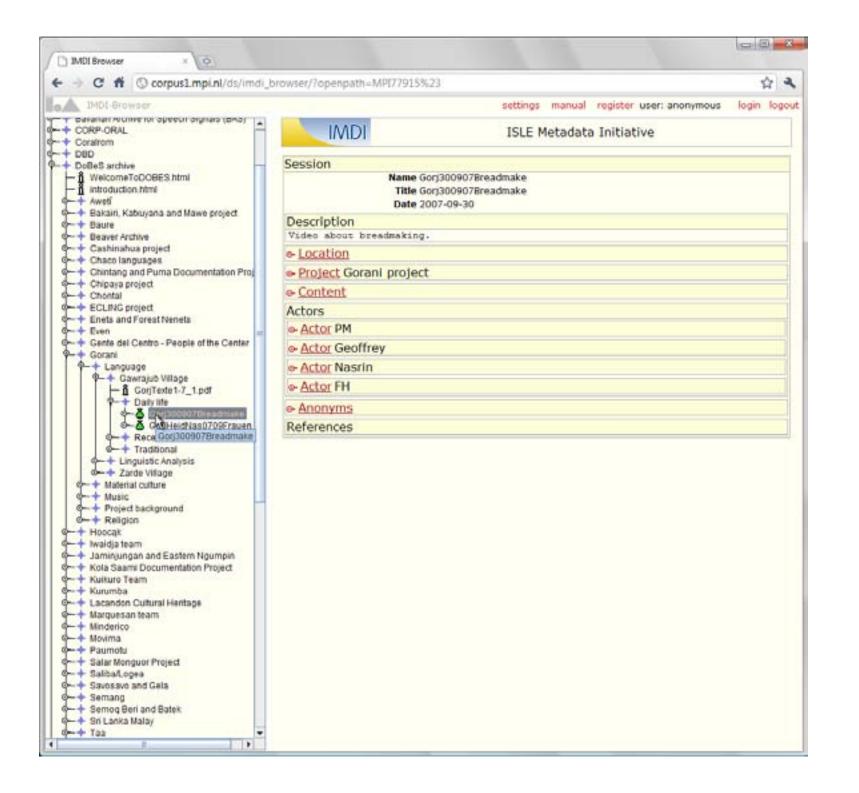
URCS - only community members are allowed access

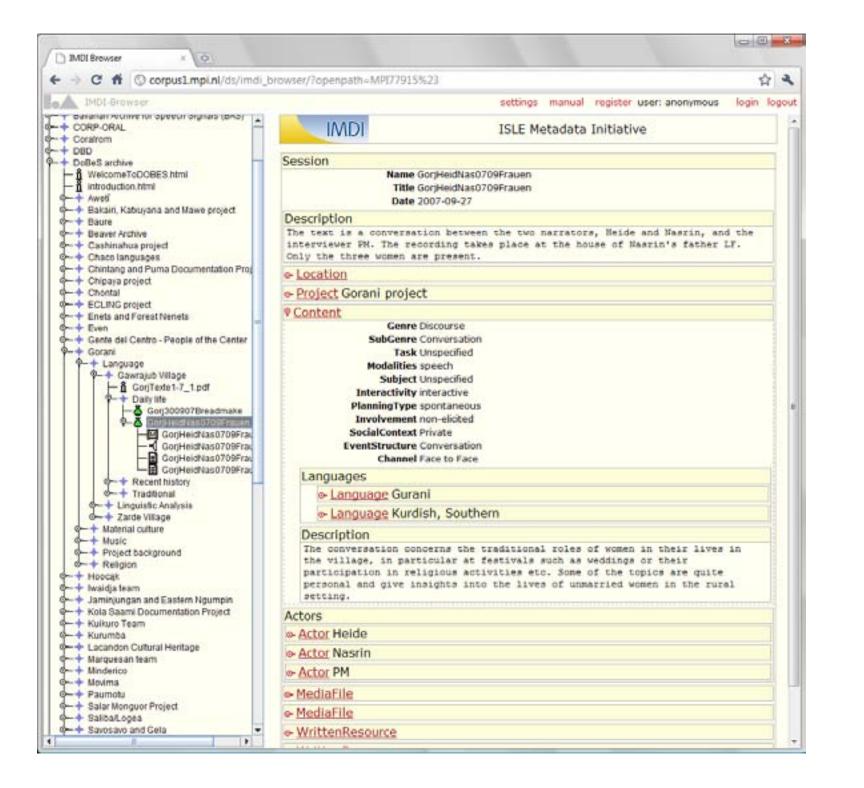
URCS - only subscribers are allowed access (requires application to depositor)

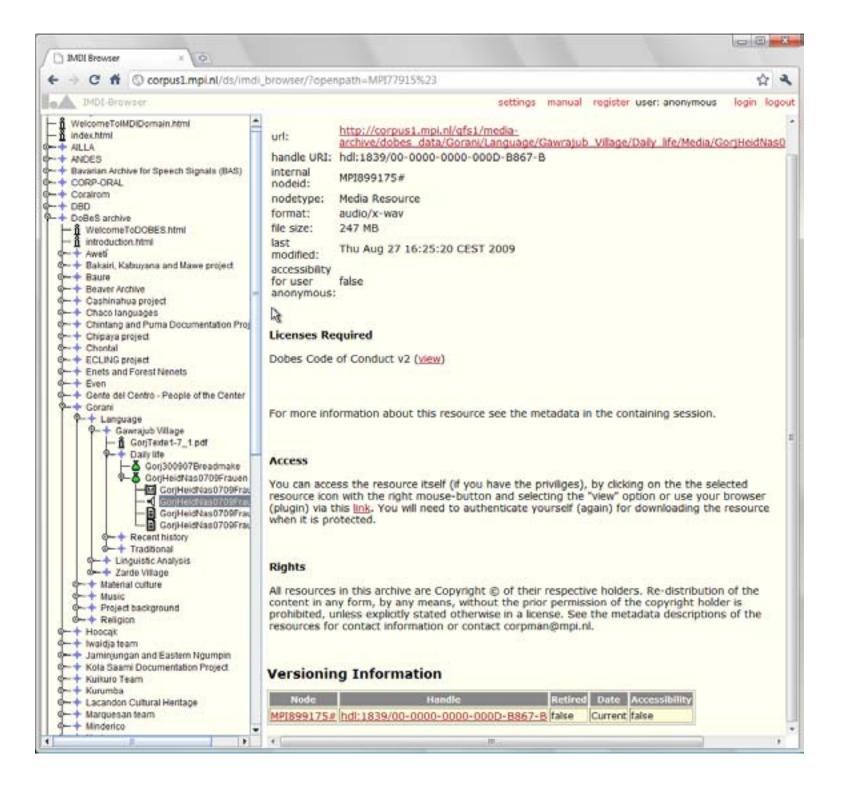


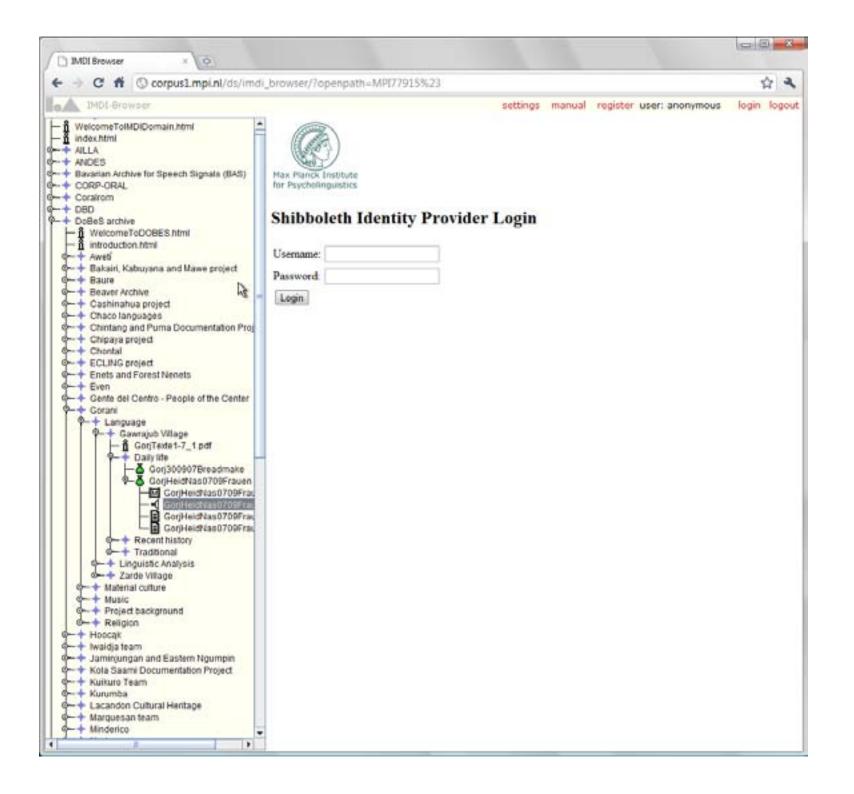
Protocol / access control

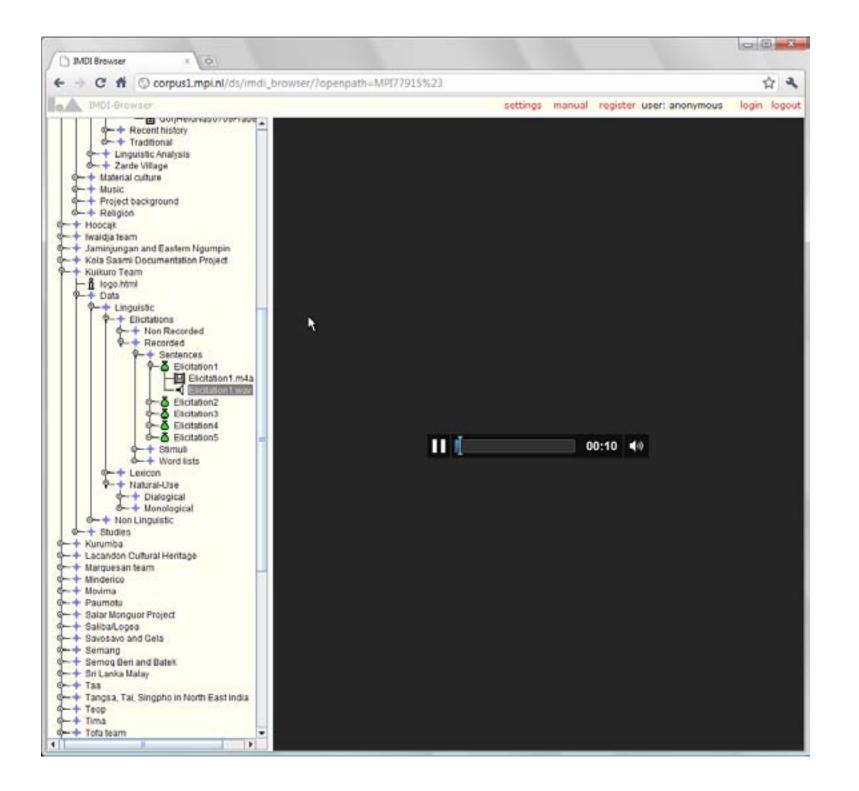


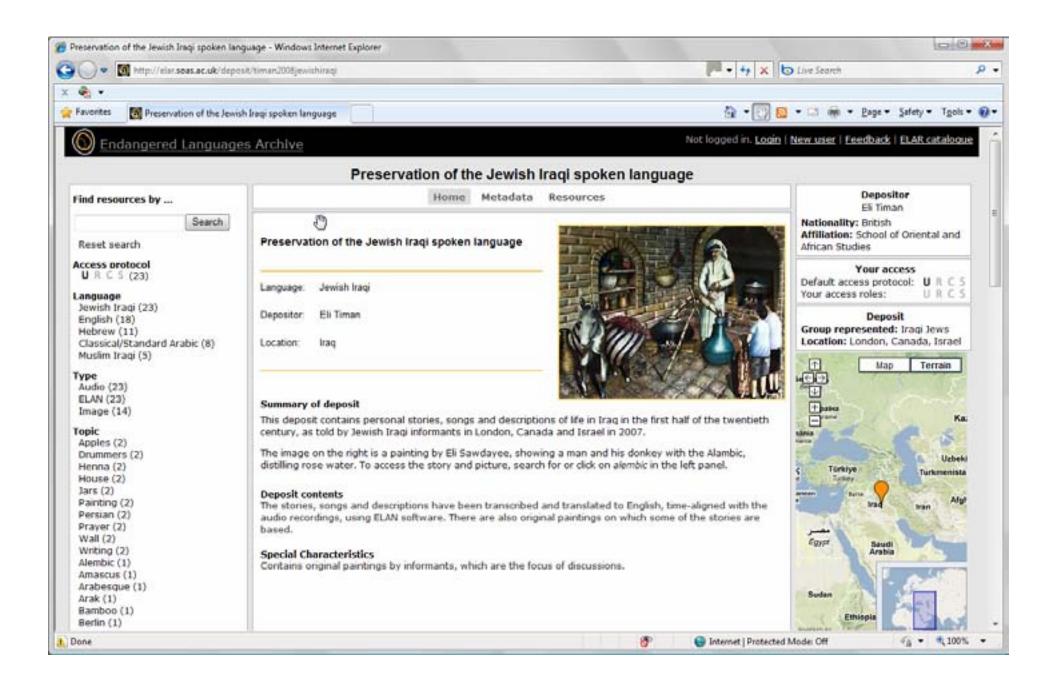


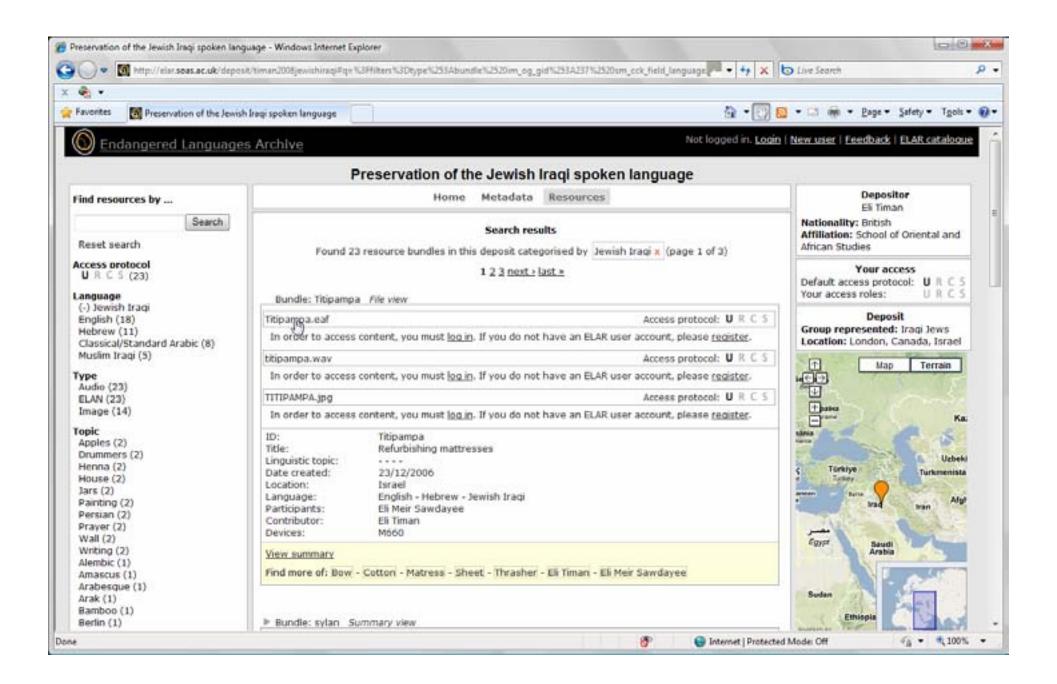


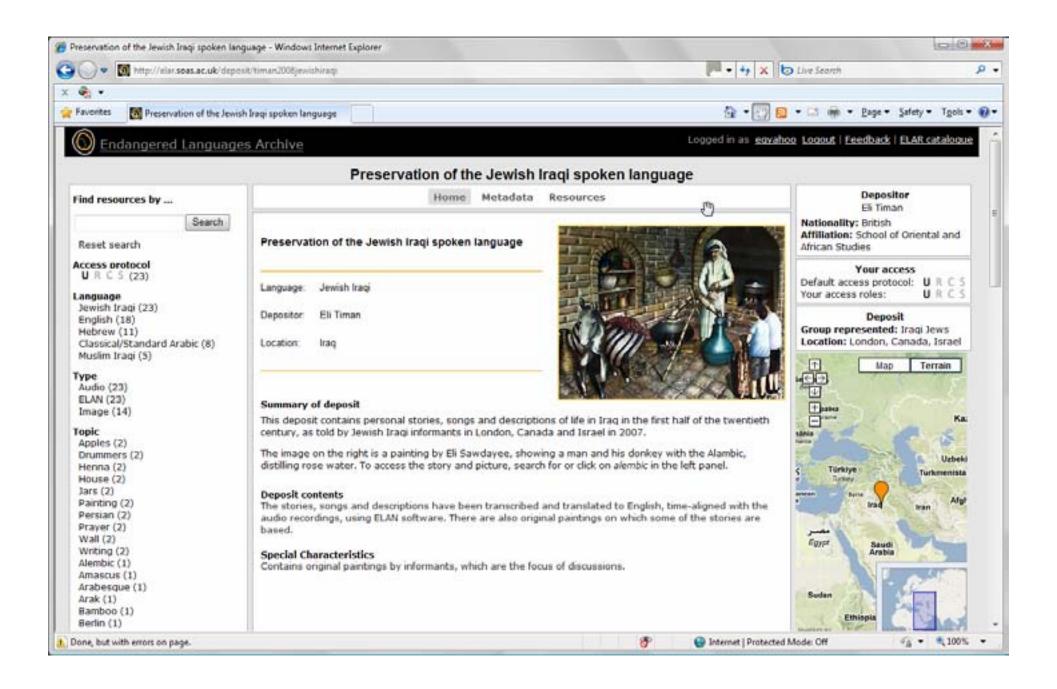


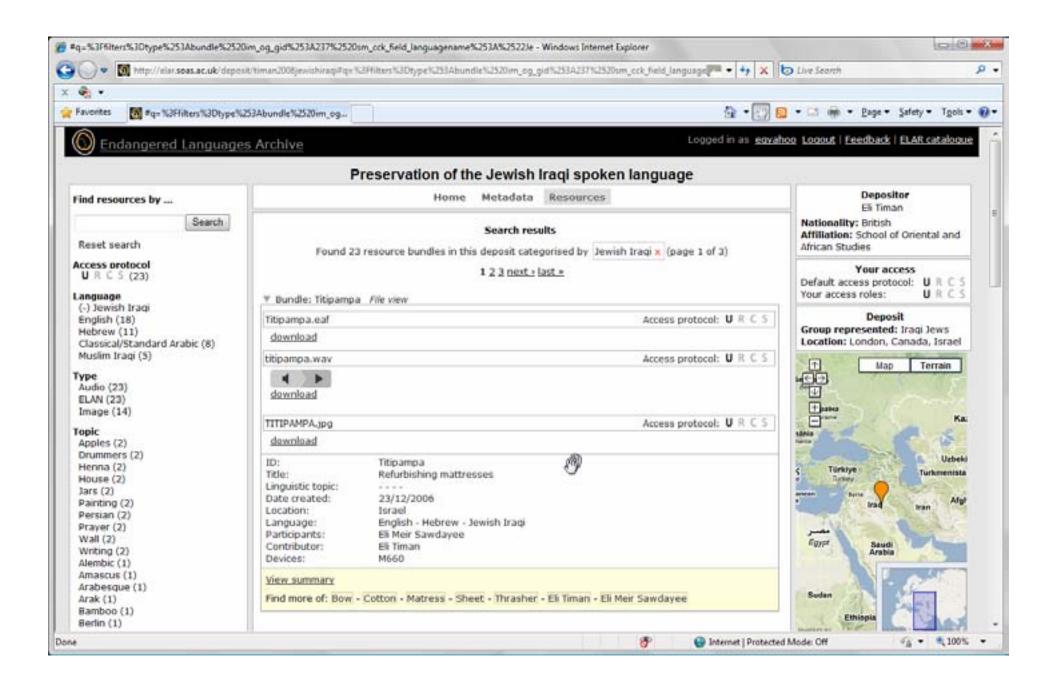


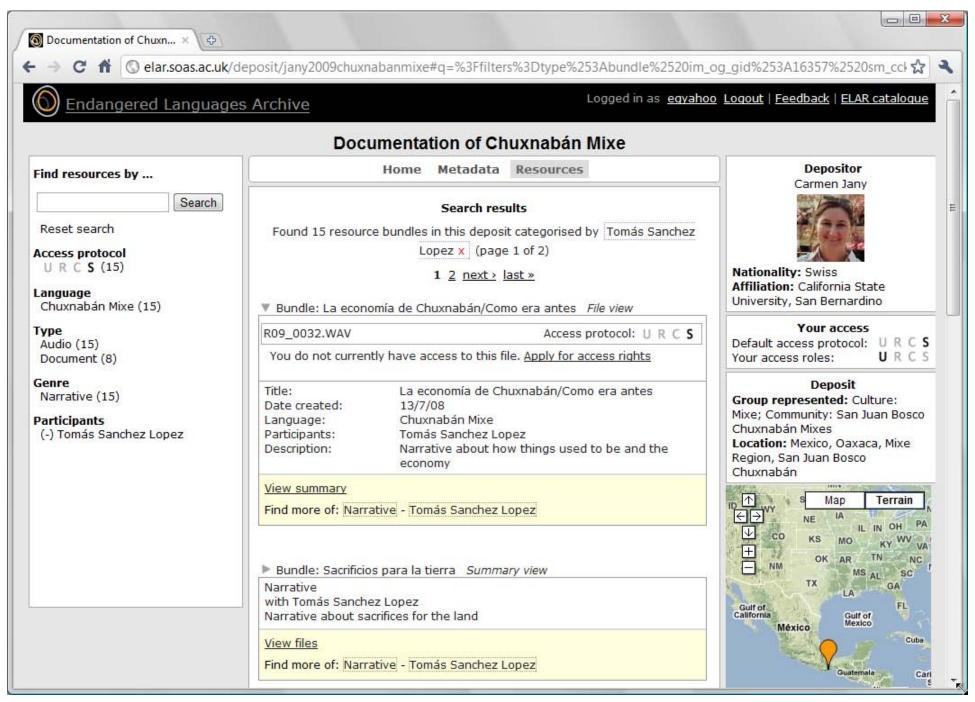






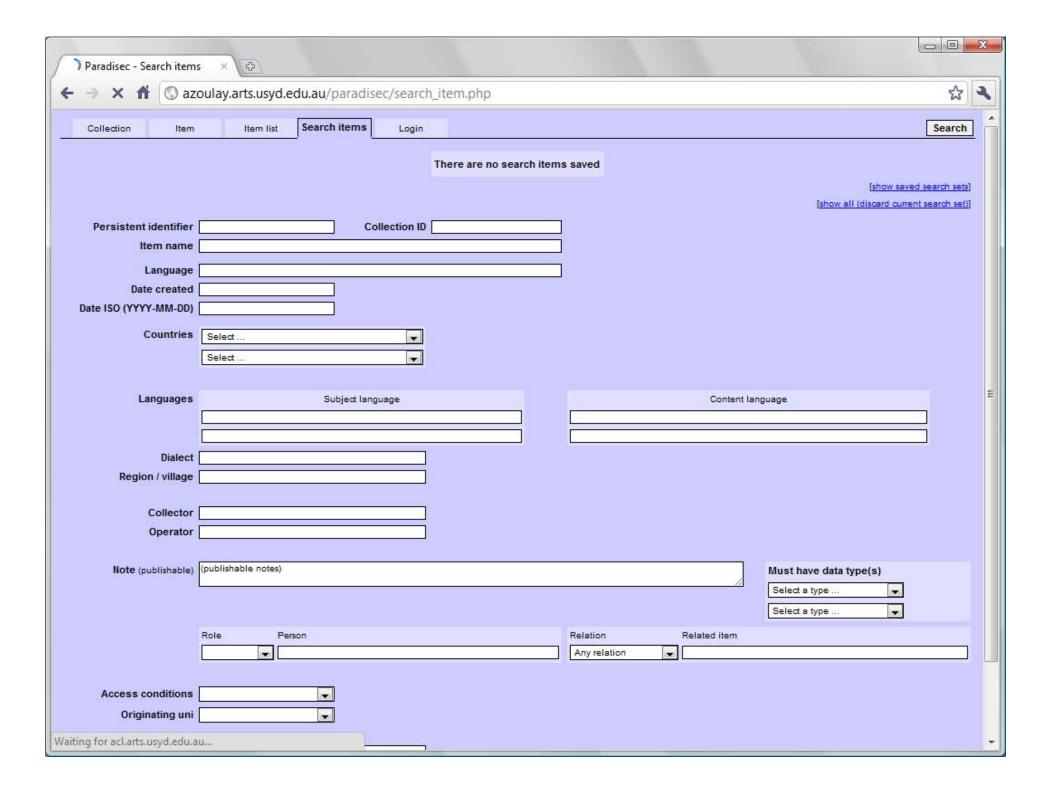


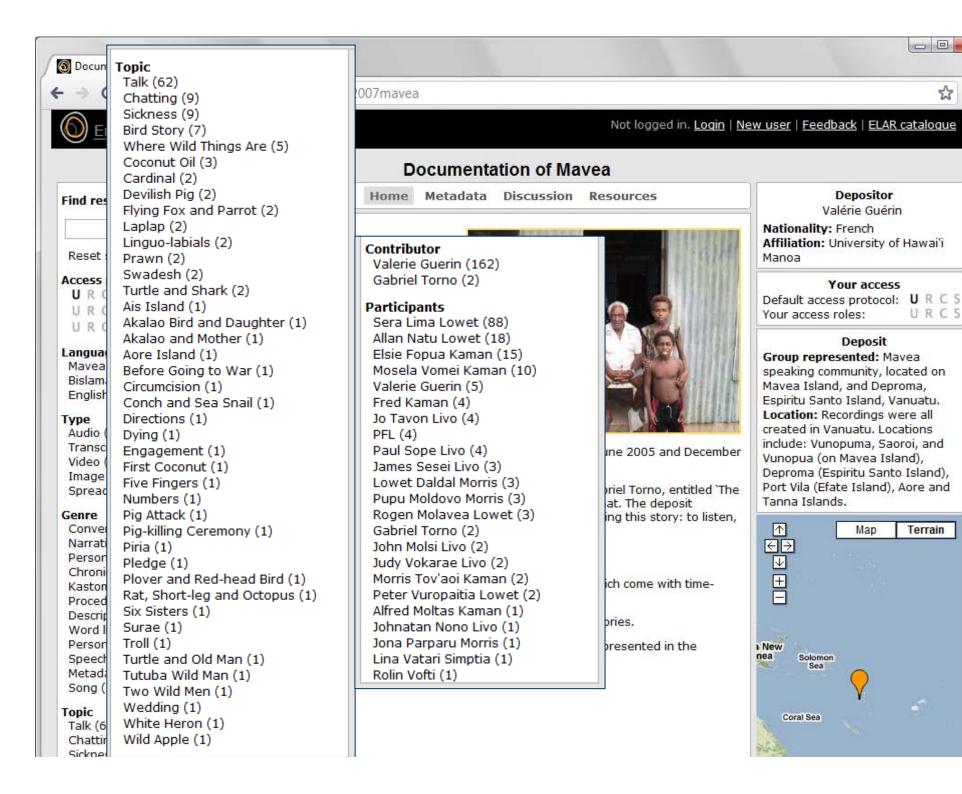






Interface: metadata





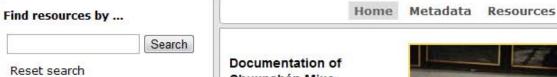
_ @ X

숪

URCS



Documentation of Chuxnabán Mixe



Access protocol

U R C S (77)

Language

Chuxnabán Mixe (77)

Type

Audio (76) Document (12)

Genre

Narrative (70) Grammar elicitation (6) Annotation (1)

Participants

Tomás Sanchez Lopez (15) Tomás Gomez Aguilar (13) Antonio Feliz Clemente (12) Juan Bautista Hernandez (10) Luisa Nolasco Martínez (7) Margarita Garcia Reyes (7) Antonio Vasquez Cecilia (6) Cuactemoc Salvador Canseco (6)

Chuxnabán Mixe

Chuxnabán Mixe Language:

[pxm]

Depositor: Carmen Jany

Mexico Location:

Summary of deposit

This deposit consists of audio recordings with transcriptions and photos of speakers of Chuxnabán Mixe, a previously undescribed language spoken in the state of Oaxaca, Mexico, resulting from fieldwork conducted between July 2008 and March 2009.

Resumen del depósito

Este depósito consiste en grabaciones de audio con transcripciones y fotos de hablantes de Chuxnabán Mixe, un idioma sin descripción previa que se habla en el estado de Oaxaca, México, que resultan de trabajo de campo realizado entre julio 2008 y marzo 2009.

Deposit contents

The deposit comprises over 70 audio recordings, transcriptions and Spanish translations of some of the narratives, and photos of the village and its surroundings. Genres include folk narratives, fables, and local history, customs and beliefs.

Contenido del depósito

El depósito consta de más de 70 grabaciones de audio, transcripciones y traducciones en castellano de algunas narraciones, y fotos del pueblo y sus alrededores. Los géneros incluyen narraciones folklóricas, fábulas, e historias, costumbres y creenciae locales





Nationality: Swiss Affiliation: California State University, San Bernardino

Your access

Default access protocol: URCS Your access roles: URCS

Deposit

Group represented: Culture: Mixe: Community: San Juan Bosco

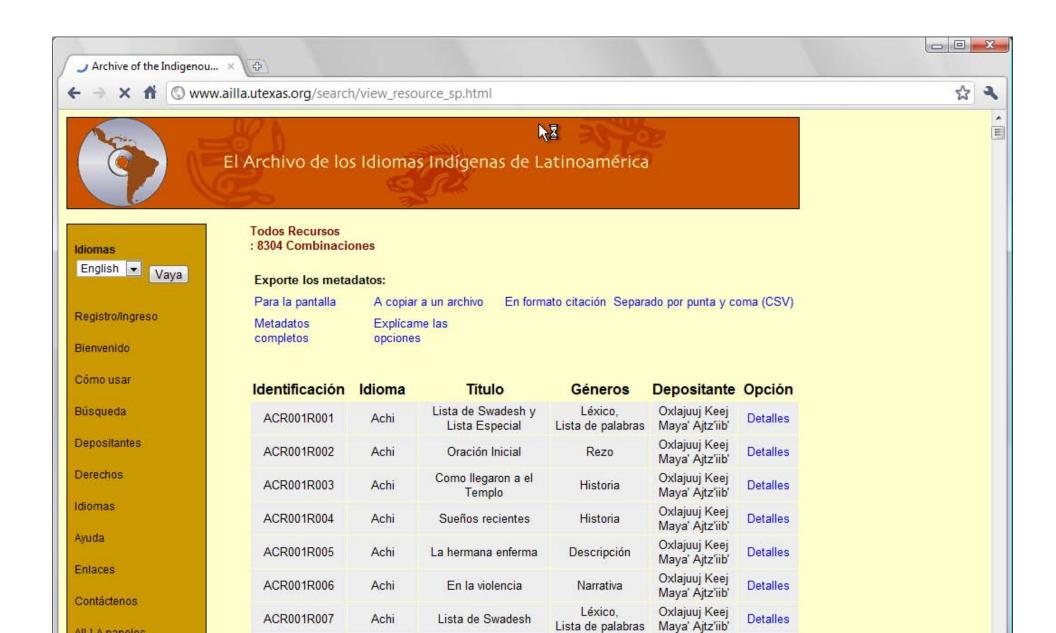
Chuxnabán Mixes

Location: Mexico, Oaxaca, Mixe

Region, San Juan Bosco

Chuxnabán





Oxlajuuj Keej

Maya' Ajtz'iib'

Oxlajuuj Keej

Maya' Ajtz'iib' Oxlajuuj Keej

Maya' Ajtz'iib'

Oxlajuuj Keej

Narrativa

Narrativa

Léxico.

Lista de palabras

Narrativa

Detalles

Detalles

Detalles

Detalles

AILLA papeles

Eventos

ACR001R008

ACR001R009

ACR001R010

ACR001R011

Achi

Achi

Achi

Achi

Historia

Lista de Swadesh

Lista Especial

Historia de su vida



What counts as access?



Access to what?

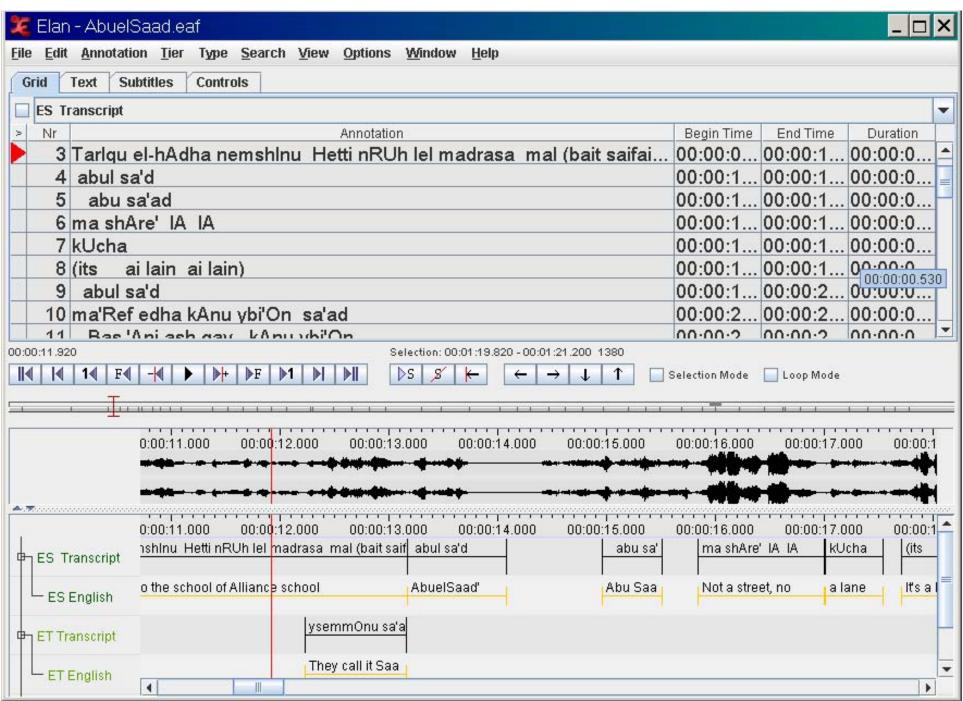
- different stakeholders want different things
 - eg descriptivists/typologists want standardised, technically encoded, linguistic system-markup
 - community members may want personally or culturally relevant items

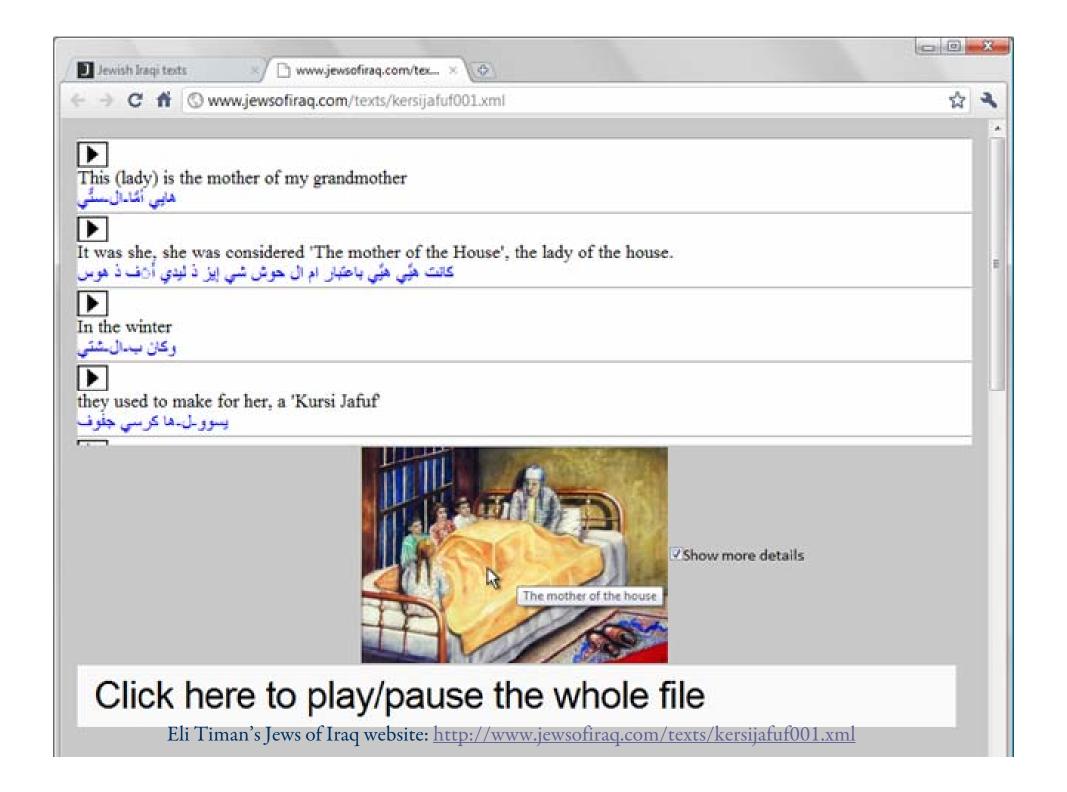


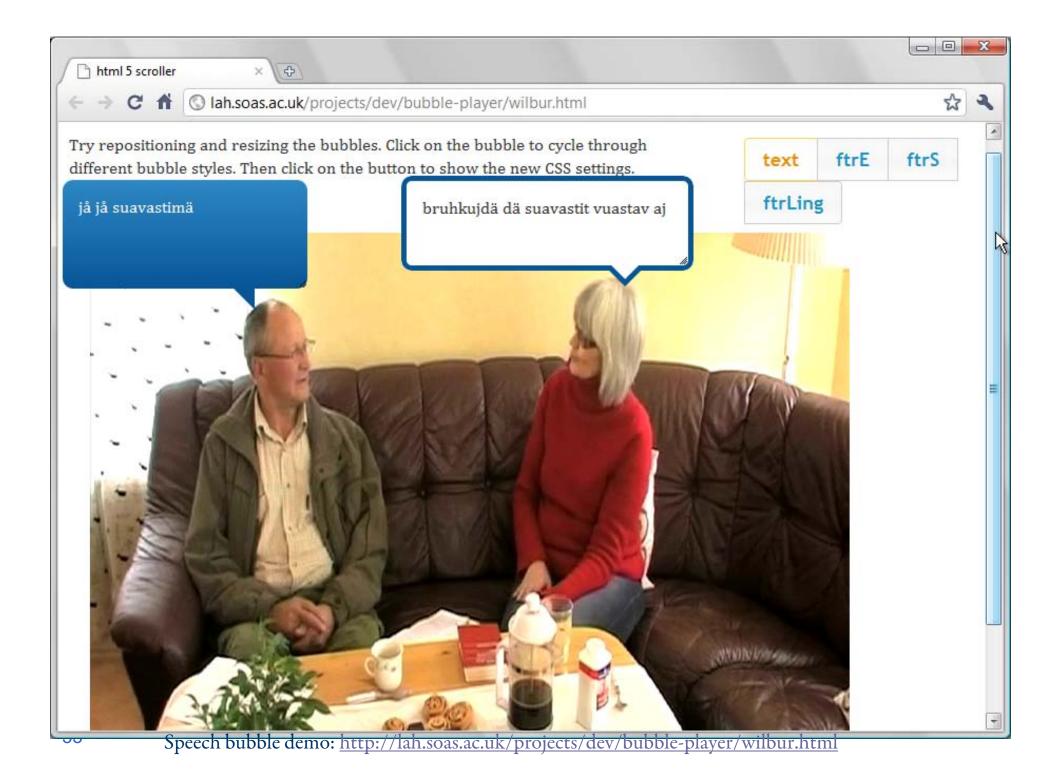
What is access?

- play file(s)?
- download file?
- make content accessible to user?











Access is also about perceptions

- perceptions of interfaces like Facebook, ELAR
 "reduces distance and power disparities, supports interactions"
- difficult access mistaken for access control
- personal appearance on website is overwhelming (NW coastal case)



Responses and conclusion

- URCS positively received
- subscription requests all agreed by depositors
- terms and conditions, onerous but accepted!
- feedback about signup process, too "academic"
- access only for registered users?
- access as access-to-upload, cf Chinese case

The next big challenge ...

documentation is the discovery and widespread use of software interfaces that make aesthetically appealing documentation materials that can be easily and flexibly used by a wide range of users, including the speech communities whose endangered languages we want to document.

- Gerrit Dimmendaal, in Language Documentation and Conservation, v4

elar.soas.ac.uk





Endangered Languages Archive

Logged in as arjukes Logout

Find a deposit:

List

Map

Help Forums

Home

Endangered Languages Archive (ELAR)

Hans Rausing Endangered Languages Project, School of Oriental and African Studies, London

ELAR preserves and disseminates digital documentation of endangered languages around the world. ELAR provides archiving for all endangered language documentation, especially the outcomes of ELDP-funded projects. ELAR also provides training and other archive-related services.

ELAR currently has about 70 deposits. Those shown below have been released for access, subject to depositors' conditions. Access conditions are shown as a graphic:

URCS - all registered users can access.

U R C S - researchers and community members are allowed access

URCS - only community members are allowed access

URCS - only subscribers are allowed access (requires application to depositor)

For other deposits, you can use the menu to the left to find deposits and see their metadata descriptions. Read more about access

Added October 2010:

Documentation of Kayardild [Erich Round]

... verbal texts, traditional food-gathering activities and elicitation from the last Kayardild speakers well versed in traditional knowledge, comprising 500 audio, video, and transcription files.



Chorote-Spanish dictionary [Nicholas Drayson]

... Diccionario Iyojwa'ja 'Lij - Kilay 'Lij (Chorote-Spanish Dictionary), including cultural notes, usage examples, scientific terms, and appendixes with outline grammar, place-names and other geographical data.

URCS Registered users of ELAR can access this deposit.

Linguistic data of Kouro [Margaret Florey]

... audio and video recorded in Rutah village, Seram Island, Maluku Province, Indonesia, including elicitation, interviews, and stories about daily life.

URCS This deposit requires you to apply to the depositor for individual access.



Documentation of Revesano [Antoine Guillaume 1

... audio and video recordings with transcriptions and photos of speakers of Revesano, a highly endangered language spoken in Reyes, Bolivia.

URCS This deposit requires you to apply to the depositor for individual access.

Added September 2010:

Documentation of Mavea 2007 [Valérie Guérin]



Preservation of the Jewish Iraqi spoken language [Eli Timan]



ja (pref) futuro

is (adv) negative

jatak'i [s] mosquito; zancu

[zjnetk'ini [adv] hasta no a

iniudc'hi vov a pegui jujtajnata (sm) une que quit

































Contact us

- about grants for language documentation
 - eldp@soas.ac.uk
- about studying at SOAS or visiting our project
 - elap@soas.ac.uk
- about the archive, and depositing
 - archive@hrelp.org
 - http://elar.soas.ac.uk
- David Nathan
 - djn@soas.ac.uk

