

La red emergente de Archivos de Lenguas Indígenas de la América del Sur

Encuentro **ALICE2** : Oficina sobre comunidades de Usuarios
Asunción, 18 noviembre 2009

Lucía Golluscio (Universidad de Buenos Aires-CONICET), Victor Miyakawa
(IIAP), Denny Moore (Museu Goeldi / MCT), Sebastian Drude (Museu Goeldi /
Universidad de Frankfurt)

Contenido

1. Información de antecedentes: Documentación lingüística y el programa DOBES
2. El archivo DOBES y la tecnología LAT
3. Los archivos regionales en América del Sur
4. La red sudamericana y planes futuros

Contenido

1. Información de antecedentes: Documentación lingüística y el programa DOBES
2. El archivo DOBES y la tecnología LAT
3. Los archivos regionales en América del Sur
4. La red sudamericana y planes futuros

Documentación lingüística y el programa DOBES

- Desde aprox. 1992 los lingüistas se dieron cuenta de que la diversidad lingüística está fuertemente amenazada
- De las aprox. 6000 lenguas en el mundo de hoy, más de 5000 pueden fácilmente desaparecer en los próximos 150 años
- Con las lenguas, se va gran parte del patrimonio cultural e intelectual tradicional de la humanidad

Documentación lingüística y el programa DOBES

- En este siglo, se crearon programas de investigación de documentación lingüística
- El objetivo principal de la D.L. es la creación de registros con datos lingüísticos y culturales
- Los datos registran las prácticas y conocimientos actuales amenazados, para futuras generaciones de hablantes, investigadores y la humanidad

Documentación lingüística y el programa DOBES

- Estos archivos contienen datos valiosísimos, que posiblemente en un futuro próximo ya no se puedan producir o registrar
- Estos archivos son (idealmente):
 - Digitales y con datos de buena calidad
 - Multi-media (con anotación & metadatos)
 - Multi-uso (científico de diversas disciplinas, prácticos)
 - sostenibles (hardware actualizado, formatos abiertos, estrategias de backup, cooperación en red e institucionalidad)

Documentación lingüística y el programa DOBES

- Un programa pionero en el área es el programa DOBES (*Documentación de Lenguas Amenazadas*), desde 2000
- Ver: <http://www.mpi.nl/DOBES>
- Hasta ahora aprox. 55 proyectos, documentando ca. 65 lenguas en todo mundo
- Aprox. 17 proyectos en América del Sur

Mapa DOBES



<http://www.mpi.nl/DOBES/dobesmap/>

Contenido

1. Información de antecedentes: Documentación lingüística y el programa DOBES
2. *El archivo DOBES y la tecnología LAT*
3. Los archivos regionales en América del Sur
4. La red sulamericana y planos futuros

El archivo DOBES y la tecnología LAT

- El centro tecnológico del programa DOBES es el Instituto para Psico-lingüística en Nimega, Holanda, Max Planck Institute
- Este instituto también es uno de los centros importantes del proyecto europeo CLARIN
- Allí está alojado el archivo DOBES: registros de audio, video, sesiones y documentos totalizando, en la actualidad 7.5 terabytes en Europa
- Actualmente, en America Latina se estima una capacidad de 4 a 8 (1 o 2 por centro) terabytes, nueva documentación esta incrementando esta capacidad de almacenamiento (se calcula 25 Tb en pocos años)

El archivo DOBES y la tecnología LAT

En el contexto del programa DOBES, se desarrollan nuevos formatos y nuevas tecnologías, como:

- Metadatos: **IMDI** (*editor, browser, www, ARBIL*)
- Annotación: **EAF** (*ELAN, ANNEX*)
- Bases de datos lexicales: **LMF** (*LEXUS*)
- Gráficos de constituyentes Sintácticos: *SYNPATHTY*
- Estructuras relacionales: *VICOS*
- Administración de archivos: **LAT** (*LAMUS: upload de sesiones, creación de estructuras etc.; AMS: administración de usuarios, derechos de acceso*)

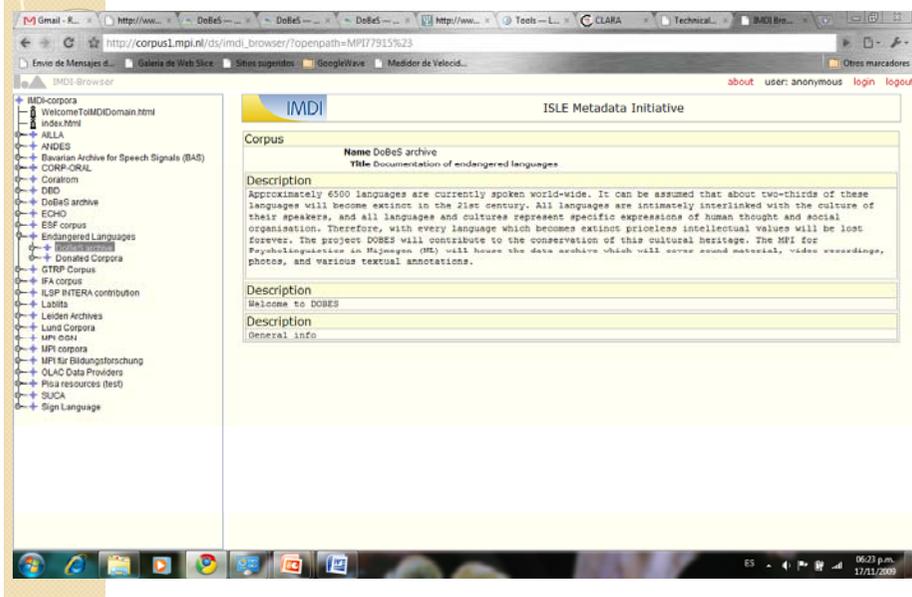
<http://www.lat-mpi.eu/tools/>

The screenshot shows a web browser window displaying the website <http://www.lat-mpi.eu/tools/>. The page features a navigation menu on the left with options like 'Inicio', 'Tools', 'Addit', 'AMS', 'Annex/Trova', 'Arbil', 'Elan', 'Imdi', 'Lamus', 'Lexus', 'Synpathy', 'Vicos', 'Papers', 'Status', and 'Related Links'. The main content area is titled 'Tools' and lists several tools with their icons and descriptions:

- Addit**: a tool to 'plug' notes to elements of the MPI archive.
- AMS**: a tool to grant and deny the access to a (part of a) corpus.
- Annex/Trova**: the Annotation Exploration tool in the MPI web-based framework for archive exploration and enrichment.
- Arbil**: Arbil is an application for arranging research material and associated metadata into a format appropriate for archiving.
- Elan**: Multimedia Annotator.
- Imdi**: create, search through and modify metadata for linguistic resources.
- Lamus**: Language Archive Management and Upload System.
- Lexus**: a web-based lexicon tool.
- Synpathy**: a tool for manual syntactical annotation.
- Vicos**: a tool that allows users to complement lexical spaces (as created by LEXUS) with ontological spaces.

At the bottom of the page, there is a section for 'últimas noticias' (latest news) with a link to 'Relacionado' (related).

Árbol IMDI



El archivo DOBES y la tecnología LAT

- El conjunto de esta tecnología se llama “Language Archive Technology” (LAT)
- <http://www.lat-mpi.eu/tools/>
- La tecnología central para la construcción y administración de archivos lingüísticos (IMDI-browser, LAMUS, AMS.) está siendo desarrollada para ser implementada en otros centros

Contenido

1. Información de antecedentes: Documentación lingüística y el programa DOBES
2. El archivo DOBES y la tecnología LAT
3. *Los archivos regionales en América del Sur*
4. La red sulamericana y planos futuros

Archivos regionales en América del Sur

- En América del Sur, hay actualmente cuatro servidores con tecnología LAT:
 - ARGENTINA: Buenos Aires, CONICET
 - PERÚ: Iquitos, IIAP
 - BRASIL: Rio de Janeiro, Museu do Índio
 - BRASIL: Belém, Museu Paraense Emílio Goeldi / Ministério de Ciência e Tecnologia
- Otros centros están surgiendo en América Latina (México y Ecuador), habiendo la posibilidad de incluir otros países.

Contenido

1. Información de antecedentes: Documentación lingüística y el programa DOBES
2. El archivo DOBES y la tecnología LAT
3. Los archivos regionales en América del Sur
4. *La red sudamericana y planes futuros*

La red sudamericana y planes futuros

Entre los centros aquí presentados, estamos construyendo una **RED de archivos**

Sulamericanos, para fortalecer:

- La interacción entre archivistas sudamericanos, investigadores y miembros de las comunidades
- El intercambio y procesamiento de datos y recursos documentarios sobre lenguas y sociedades indígenas
- El desarrollo de nuevas actividades colaborativas **interdisciplinarias** (ej. articulación entre recursos de información sobre Ciencias Sociales y Ciencias Naturales)

La red sudamericana y planes futuros

- El objetivo práctico y tecnológico inmediato es el intercambio de datos entre los centros para facilitar acceso mutuo y procesos (*backup*, edición colaborativa, entre otros)
- Junto con el MPI Nimega se están creando soluciones para el intercambio en el nivel lógico (acceso unificado, intercambio federado, identificación compartida etc.)

La red sudamericana y planes futuros

- La visión es consultar y compartir los datos en una red tipo grid, entre los SA archivos y con centros y redes internacionales, como DELAMAN y la red CLARIN (Europa)
- Para esto se necesita:
 - Mayor la velocidad (banda ancha) de la conexión entre los centros e internacionalmente
 - Crear mecanismos de intercambio automáticos y procesos como (*backup*, edición colaborativa, etc.)
 - Expandir la capacidad de almacenamiento en cada centro

- www.mpi.nl/DOBES/
- www.hrelp.org
- www.ailla.utexas.org
- www.delaman.org
- www.caicyt.gov.ar/DILA
- www.museudoindio.gov.br
- www.museu-goeldi.br
- www.iiap.org.pe

La red emergente de **Archivos de Lenguas Indígenas de la América del Sur**

*Encuentro **ALICE2** : Oficina sobre comunidades de Usuarios*
Asunción, 18 noviembre 2009

Lucía Golluscio (Universidad de Buenos Aires-CONICET), Victor Miyakawa (IIAP), Denny Moore (Museu Goeldi / MCT), Sebastian Drude (Museu Goeldi / Universidad de Frankfurt)