LAMUS: Language Archive Management and Upload System

Alexander König
The Language Archive - Max Planck Institute for Psycholinguistics
Nijmegen, The Netherlands
MPI Archive Data

- Large increase in archivable data (currently ~ 50 Terabyte, almost 200,000 sessions, ~ 500,000 resources).
- MPI archive is open for external donators

How to keep control over this amount of data?
Language Archive Management and Upload System: Researchers can build corpus/upload resources. LAMUS checks formats, consistency (gatekeeper)
LAMUS features

- Web-based operation.
- Workspace principle (isolated work environment).
- Create and modify the archive's IMDI corpus structure.
- Replace existing metadata and resources in the archive.
- Upload new metadata and resources into the archive.
- Gatekeeper function: not all data formats are accepted to keep archive consistent and future proof.
- Persistent identifiers, Versioning system
Archive Organization: IMDI domain

C  Corpus node: defines structure, linked session nodes.

S  Session node: detailed metadata description, linked resources.

m T  Resource: text, multimedia.
MPI archive organization

Interaction between researcher and archive: browsing/search/visualization

(IMDI) browser

Web apps.

ARCHIVE

metadata

annotations media files

LAMUS
- create workspace
- upload files
- create links
Work Flow Overview

- create workspace
- upload files
- create links
- submit workspace
Work Flow Overview

- Open LAMUS in web browser: http://corpus1.mpi.nl/jkc/lamus
- 1st time: ask corpus manager for write rights.
- Select your workspace (as low in tree hierarchy as possible!)
- Upload metadata, resources, change corpus structure, etc.
- Submit the workspace, apply changes to actual archive.

Optional:
- Manage access rights to the resources using AMS (metadata is always accessible for everyone).
Register as a user

Welcome to LAMUS

New users

- Register as a new user

Users with a LAMUS account

- Create new workspace
  Select an archive node to create a new workspace

- Select existing workspace
  Select one of my existing workspace(s)

- Request storage space
  Send a request for storage space to upload audio/video resources

LAMUS Administrators

- Management Console
  Manage ingest requests and system state

LAMUS Information

- About LAMUS
  LAMUS version and preferences info

Register New LAMUS User

Contact information

- Full Name: Janne Groot
- Postal Address: Wundthaan 1 65293 ID Nijmegen
- Affiliation: MPI
- Email Address: JGroot@nethunt.com

Archive use

Describe your existing relation with this archive or your purpose to use the archive:
I want to donate some movies showing traditional fishing techniques used by the Kamoso people of S.Y. Asia.
Destination: http://corpus1.mp1.nl/gst/Species/Parrotfish/Parrotfish/JGroot.sact

If you need to have access to an existing corpus, please state its name or find it with the archive browser.
If you want to upload a completely new corpus, specify a short name for it.

Corpus name: KamloopsGroot

Account Login Info

If you have already an user account at the Max-Planck Institute for Psycholinguistics, please use your existing username. If you do not have an existing account, give your (new) user name and an initial password.

- User ID: JGroot
- User Password: ********

Submit

Thank you for your registration.
You will get a confirmation email from corpus management soon.
Select a new workspace

Welcome to LAMUS

New users
- Register as a new user
- Register as a new user to get access to LAMUS

Users with a LAMUS account
- Create new workspace
- Select an archive node to create a new workspace

Open existing Workspace

Welcome to LAMUS

New users
- Register as a new user
- Register as a new user to get access to LAMUS

Users with a LAMUS account
- Create new workspace
- Select an archive node to create a new workspace
- Select existing workspace
- Select one of my existing workspace(s)

Ingest Request Choice

Please choose one of the following existing workspace(s):

- Starting Date: 2005-08-03, Last accessed at 2005-08-09, IngestRequestID: 1, TopNode: file
- Starting Date: 2005-08-03, Last accessed at 2005-08-09, IngestRequestID: 2, TopNode: file
Manage the workspace

Right mouse button on a node: context sensitive tree menu.

Options like:

- Adding corpus nodes, changing structure
- Renaming
- (Un)linking session nodes, resources
Example: add a new resource

Step 1: upload the local resource to the workspace.

After upload the file types and formats are checked, only accepted ones will be placed in the “Unlinked files” container.
Example: link resource

Step 2: Link the resource to a session node.

Choose file type.
Example: link resource

Step 3: add additional information and submit

The resource should now be linked to the session in the corpus tree of the workspace.
Example: submit workspace

Step 4: Submit the workspace into the actual archive
LAMUS makes it possible to:

1. Allow researchers to manage their own corpora.
2. Check for consistency.
3. Update databases and indexes (used by search engines).
Be your own corpus manager!

More information at:

http://www.lat-mpi.eu/

➔ Other tools (ANNEX, LEXUS, etc.)
➔ Manuals
➔ Fora