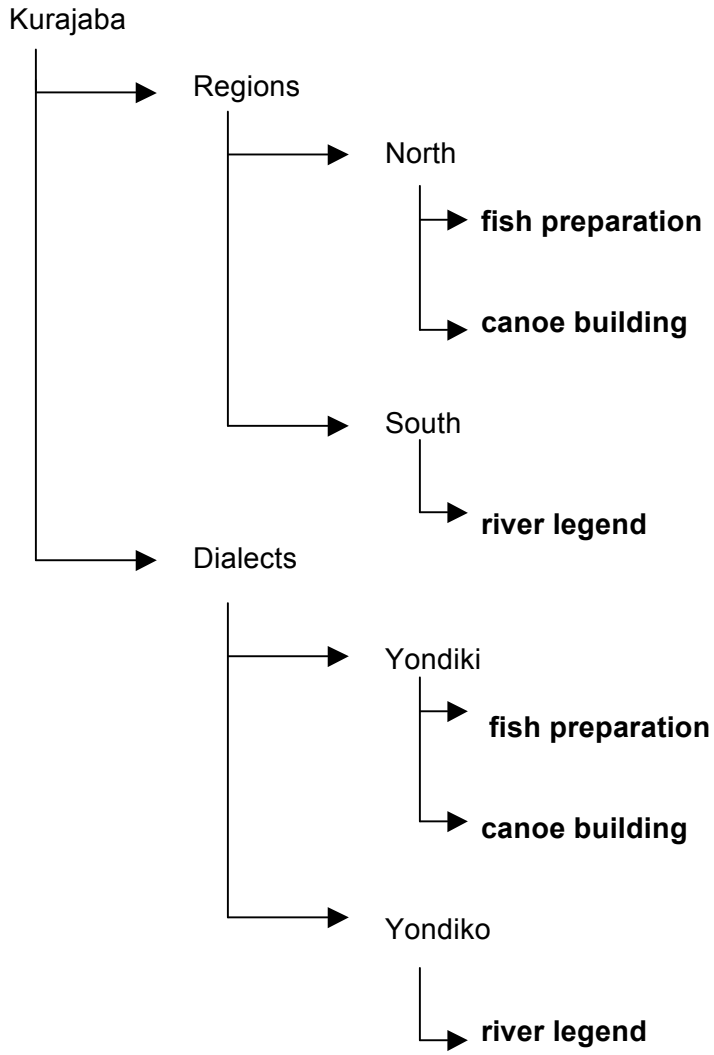
# LAMUS: Creating a corpus structure

**Assignment**
Step by step, create a corpus structure as described below with LAMUS.
Reference: A4 guide LAMUS

**Information about the corpus to be described with the metadata:**
The corpus consists of language data for the fantasy language Yondik, spoken in Kurajaba (a fantasy land in Asia). Yondik has two dialects: Yondiky spoken in the North of the country and Yondiko spoken in the South. The Northern part of the country borders the Ocean and the people live mainly of fishing, the southern part is mountainous with many large rivers. The basic corpus will look like this (in **bold** are the **session nodes** containing the imdi metadata). In this exercise you will create the corpus structure using the Language Archive Management and Upload System: LAMUS.



**Usernames and Passwords**
For the use of LAMUS writing rights are required on the node where you want to add files.
For username and password please refer to the one you have been handed.
In the Browsable corpus we created top nodes, group01, group02, group03....group10
The groups are situated under MPI-corpora --> DEMO --> DOBES training.
User dobes01 has read and write rights on node group01 etc.

**Data:** the metadata files and resources that will be used in this exercise are in the folder called "lamus_files".

1. Create a workspace
Go to the LAMUS home page on http://corpus1.mpi.nl/jkc/lamus and create a workspace at the top node of your group. It takes a moment before your workspace is initialized. When ready you will see the LAMUS interface with your top node on the left and LAMUS-help in the main window.

2. Create a corpus structure/linking corpus nodes
Add a Kurajaba corpus node to your top node and add two corpus nodes to this Kurajaba node: Regions and Dialects. Add two corpus nodes to Regions: North and South. Finally add 2 corpus nodes to Dialects: Yondiki and Yondiko. Your basic structure is ready.

3. Adding descriptions to the corpus nodes
In some cases you might wish to add a description to your corpus node.
Add some information to the Kurajaba node (e.g. some project information).
Check if the descriptions are correct through the option 'view node' which permits to see into the corpus imdi file created by LAMUS.

4. Uploading session nodes
Before you can link your session nodes to your corpus, you need to upload them into your LAMUS workspace. Upload the three session nodes from the lamus_files folder: f*ish preparation*, *canoe building* and *river legend*. Check in the 'unlinked files' menu that your upload was successful.

5. Linking the session nodes
Link the three sessions as depicted on the previous page, but ***only*** to the North and South nodes in Regions.

6. Saving your workspace
When you want to quit working on your workspace, but you have not finished the corpus, you may save your workspace. Click *Save and Logout* to save your workspace. To re-open your workspace, go to the LAMUS main menu, choose *Select Existing Workspace* and select the workspace from the list.

7. Create parallel structures
In the corpus design you can see that the session/imdi files are linked in two locations. Under Region the session files are already linked in. It is not necessary to upload these imdi files again; you can create an extra link to the same imdi file. To do this, click on the node where you want to make the link (Yondiki, under Dialects), press the control button and click on the session file that you want to link to Yondiki. Now both are selected. Right click and select: 'link selected nodes'. Also create a link from Yondiki to canoe building and a link from Yondiko to river legend.

8. Linking Resource files
The resource files need to be linked to the session file.
First upload the resource files from zip file (one *.mpg file and one *.eaf file) and check in Unlinked Files that your upload was successful. Next link the video file to the one of the sessions and also link the eaf file to that session. Note that the mpg file is a media file and the eaf file is a written resource. Please note that the media and written resource will also be approachable from the parallel linked session file.

9. Submitting your workspace

When the work on the corpus is ready and you have checked that everything is correct, you may submit your workspace. Click S*ubmit Workspace* in the Lamus menu. After a while (time depends on the size of the corpus, but varies from 1 minute to 20 minutes) you will receive an e-mail that your corpus was successfully ingested in the corpus. The mail will be send to the email address that was assigned to user group0x in the case of the training, but in the real world it will be sent to your email address if you have your own LAMUS account. Go to: http://corpus1.mpi.nl In the browser go to the MPI corpora --> DEMO --> DOBES training --> group(01..10) and check if your corpus is correct.

**! Note:** LAMUS will only accept one workspace on a node at the same time. If you try to create a workspace on a node which already has a workspace by someone else or by yourself, LAMUS will refuse to create a new one. Sometimes you can already see your newly updated data in the browser, but still LAMUS will not have liberated the node for new workspaces. LAMUS will still be busy calculating the access rights for the new files. Always wait for the mail before starting a new workspace!!

## Arbil Roundtrip

1. Open Arbil. Refresh the Remote Corpora area to be sure you have the newest changes.
2. Import the Kurajaba corpus you have created in the exercise above into your Local Corpus area.
3. The researcher 'Alexander Miller' who created this corpus has made a decision to alter his life is now called 'Alexandra Miller'. Change his name in all three sessions. Note that s/he appears both in the project node and as an Actor.
4. When the corpus was created, Alex forgot to add contact information. Add his/her e-mail and postal address to the project node in all three sessions.
5. Save your changes and export the Kurajaba corpus
6. Log in to LAMUS and create a new workspace on your group node.
7. Upload your Arbil export in LAMUS (you can select the whole folder)
8. Now use the 'replace tree' function to update the Kurajaba corpus node.
9. Submit the workspace
10. Go to the IMDI browser to see the effect of the change.