

Working with CMDI in Arbil

Jeroen Geerts - September 2016

The Language Archive has been migrated to the CMDI metadata standard. CMDI metadata is based on profiles; each containing a certain amount of components. The TLA has currently published 2 general metadata-profiles and 1 specific profile for Sign language. They are:

- Lat-corpus
- Lat-session
- Lat-SL-session

The migration to CMDI requires a somewhat different workflow approach with regards to Arbil. You should already be familiar with using Arbil in IMDI. Otherwise consult the manual first. This guide will cover the following points:

1. Updating the remote corpora in Arbil
2. Setup & add metadata-profiles to Arbil
3. Workflow overview
4. Converting IMDI to CMDI
5. Update metadata using CMDI

All software and manuals referred to in this guide can be found at: <https://tla.mpi.nl/tools/tla-tools/>

1. Updating the remote corpora in Arbil

After the migration, the remote IMDI corpus locations have become outdated, and these need to be updated. To do so, open Arbil and make sure you have the latest version (go to *'Help>Check for updates'*). Then highlight all the remote locations by using shift-click, then right-click and select 'Remove remote location' (fig. 1). Alternatively, you can remove a location one by one, by right-clicking each individual location and selecting 'Remove remote location'; this may be useful in case you already have some remote CMDI locations in the list.

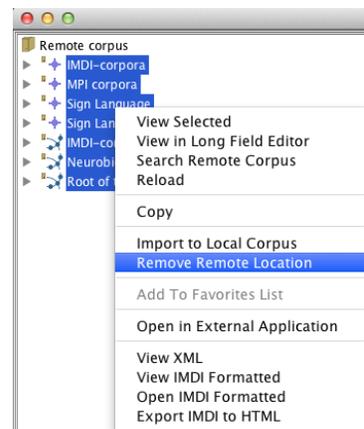


Figure 1: Remove remote locations

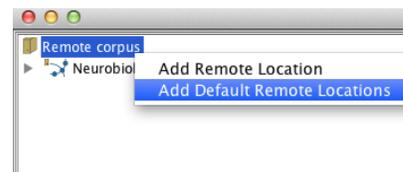


Figure 2: Add remote locations

Next, you will need to add the updated locations. Right-click the 'Remote corpus' node and select 'Add default remote locations' (fig. 2). The locations will be added and you can continue working with Arbil.

Note: You can add a custom remote location by using its handle url. You can get these from the Archive using the 'Bookmark' button for a certain node and removing the '@view' part from it. To do so, right-click 'Remote corpus'>'Add remote location' and paste the handle url you would like to add, for example: <https://hdl.handle.net/1839/00-0000-0000-000F-12AE-0>

2. Setup & add profiles to Arbil

To be able to work with CMDI and the lat-profiles, they need to be added to Arbil.

If you are running Arbil for the first time, select both 'IMDI' & 'CMDI' as the metadata standard from the setup wizard. Choose both 'lat-corpus' and 'lat-session' profiles (lat-SL-session is only used in the Sign Language department) from the list of profiles (fig. 5).

In case you have been using Arbil with the IMDI standard, you can re-run the wizard. From the main screen, go to 'help>run setup wizard'. Next, select both 'IMDI' and 'CMDI' as the standard. Select both lat-profiles from the next screen (fig. 5). Note that you can only use either IMDI or CMDI components inside a project.

When you already have configured Arbil to work with CMDI, go to 'options' from the menu and select 'Templates & profiles'. In the dialog that appears, tick the two profiles from the list on the right-hand side (fig. 5). Next, go to 'help>run setup wizard' and select both 'IMDI' and 'CMDI' as the standard.

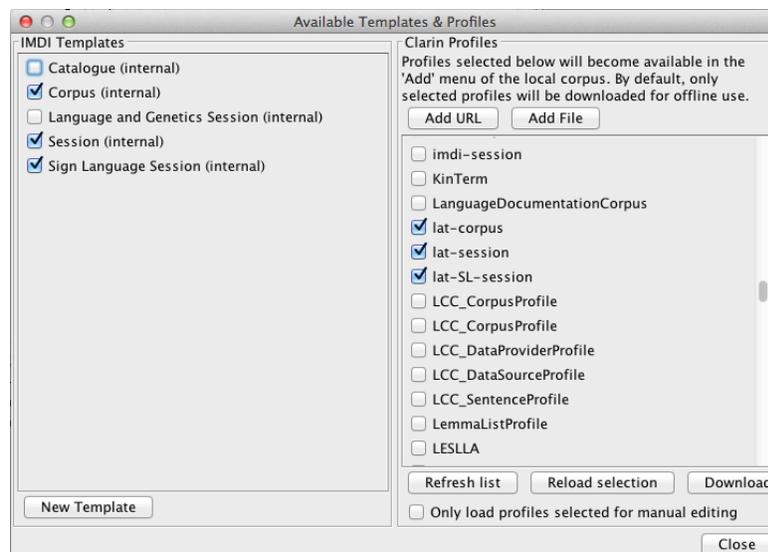
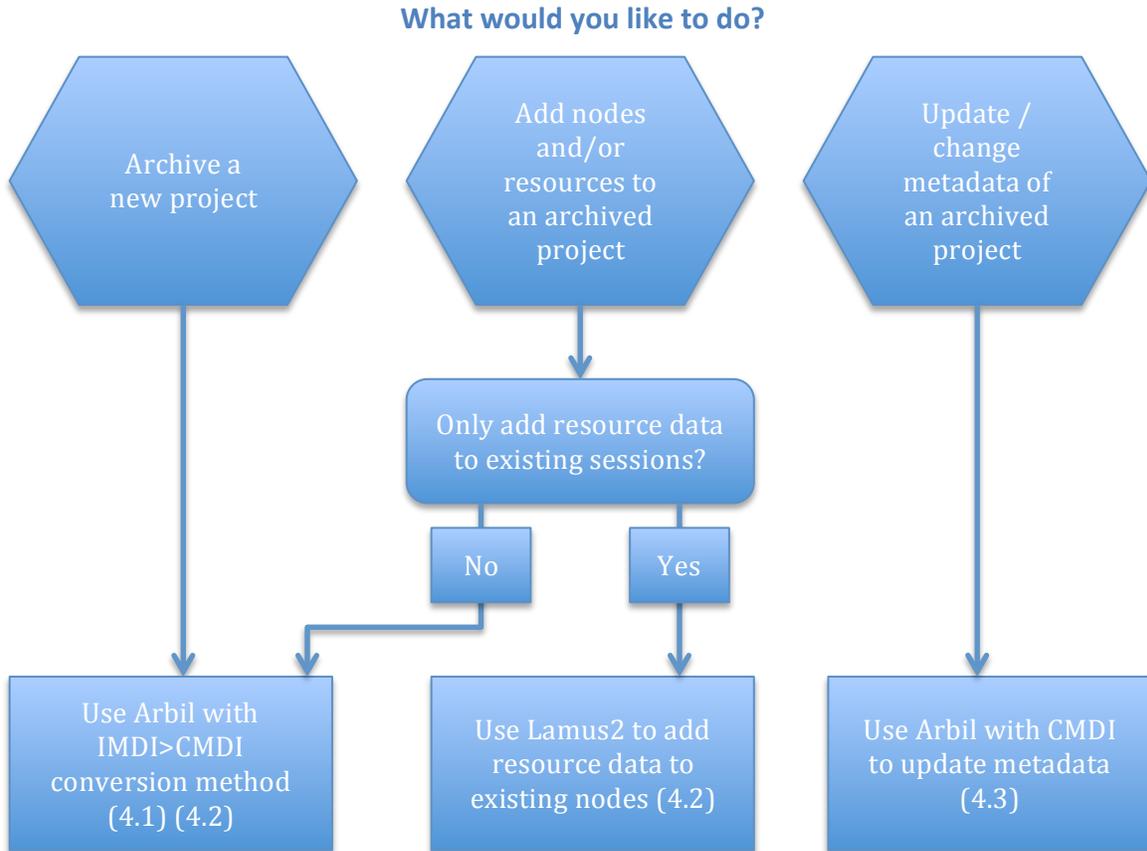


Figure 5: CMDI profiles dialog

3. Workflow overview



This flowchart should help you decide what approach you will need to take in order to archive or update resources and/or metadata successfully. Each method is described in more detail on the following pages.

IMPORTANT: THE MIGRATED CMDI ARCHIVE DOES NOT ALLOW ARCHIVING OF EMPTY SESSION NODES. THIS MEANS THAT YOU CAN STILL CREATE A PROJECT IN ARBIL CONTAINING EMPTY SESSIONS AND EXPORT IT. HOWEVER, ONCE YOU UPLOAD THAT PROJECT IN LAMUS2, YOU MUST LINK AT LEAST ONE RESOURCE TO THE SESSION, OTHERWISE THE INGEST WILL BE REJECTED BY THE SYSTEM.

4. Converting IMDI to CMDI

The latest version of Arbil is also able to convert IMDI based projects to CMDI metadata. This means that your previously created (but not yet exported) projects in Arbil can still be used and exported for use with the new CMDI archive.

1. Create new projects using IMDI

It is possible to create a new project with Arbil using the old IMDI workflow and exporting it as CMDI. For many users, this will be the quickest way to archive a new project, we recommend that you take this approach.

Be aware that you cannot import a remote corpus as a starting point with this method anymore, as the remote corpora are now all CMDI-based. Instead, just create an empty IMDI corpus- or session-node in the Local corpus, name it and start your project from there.

- Conversion to CMDI

The conversion from IMDI>CMDI can be done when exporting; to do so right-click the desired node and select 'Export' (fig. 3). Next, click the 'more options' button in the dialog and check 'convert IMDI to CMDI' (fig. 4).

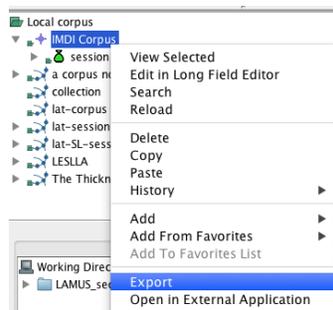


Figure 3: Export IMDI node



Figure 4: Convert IMDI to CMDI selected

2. Add nodes/resources to an archived project

In case you have already archived a project and you want to add any additional corpus- or session-nodes containing (meta)-data, you can use the above conversion method as well. Create your IMDI corpus or session, fill out the metadata, add resources and export it using the conversion method. Once again, you will have to start without importing the remote corpus from which you would like to work on.

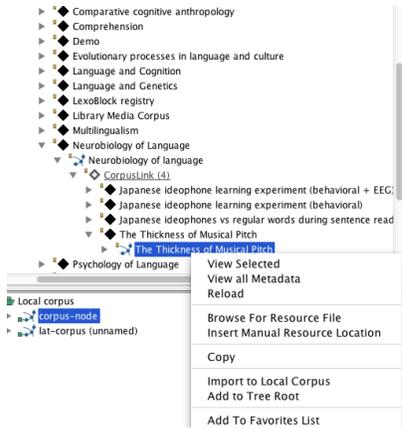
In case you only want to add resources to an existing node in the archive, use Lamus2 to upload and link these resources. Refer to the Lamus2 manual for detailed information.

3. Update metadata for archived project nodes

If you want to change or update metadata in an existing archive, you will need to work with Arbil using CMDI. Chapter 5 of this guide will explain the steps to take to update metadata for CMDI based structures within Arbil. Any additional resources can be added later with Lamus2.

5. Update metadata using CMDI

To update metadata from an archived project, first import the desired node from the remote corpus to your local corpus (fig. 6). Make sure that remote location have been updated (see section 1).



Warning: the black squares \blacklozenge indicate a component, in this case a corpus-link. To successfully import a single top-node, expand the corpus-link component and right-click the corpus-node represented with the 3 dots (see fig. 6). Otherwise, all nodes from Neurobiology of language (in this case) will be imported.

Corpus- and session nodes in CMDI are always displayed with the three dots icon.

Figure 6: import node to local corpus

Once the node is in your local corpus, you can expand it and locate the node containing the metadata you would like to update. Right-click on the component and choose 'view selected'. From there, you are able to change the value of the chosen component (fig. 7).

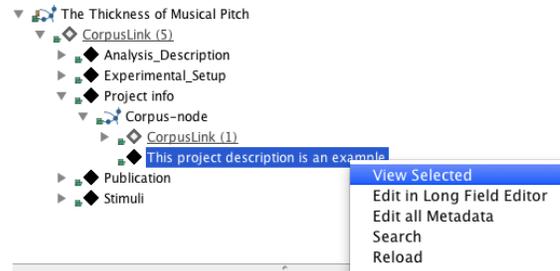


Figure 7: view and change a metadata value

You can also add additional metadata like extra actors, for instance. To do so, right-click the proper component (in this case 'actors') and choose the proper sub-component (fig.8).

Right-clicking a corpus- or session-node (represented with the 3 dots) and choosing 'add>' will show a complete list of possible components available for that node. Using this method, you can add new components to your previously archived metadata.

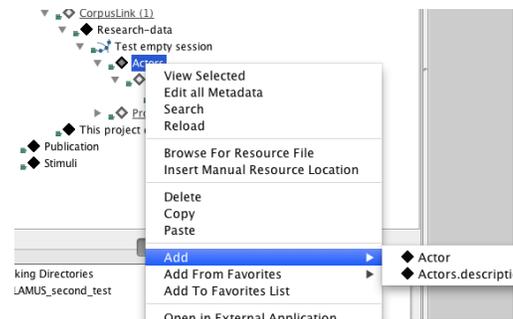


Figure 8: add an additional component

When done, export the updated CMDI node(s) and use Lamus2 to upload the metadata. In this case, you can use the 'Replace' function. Please refer to the Lamus2 manual for more details.

Adding resources to updated metadata

If you want to add additional resources to the updated metadata, you can do so with Lamus2. First update the metadata, export the node(s) and upload it with Lamus2. Replace the existing node with the updated one, then upload the resources and link them to the node. Finally, submit the updated archive.