Arbil

Arbil 2.6, for editing and managing CMDI metadata

This manual was last updated in November 2015

The latest version can be found at: http://http://tla.mpi.nl/tools/tla-tools/arbil/

Anastasios Vogiatzis (June 2012)

Claudia Zabeo (December 2012)

Janna Voelpel (September 2013)

Francesca Bechis (July 2014)

Steffen Kaiser (November 2015)

C L A R I N
Common Language Resources and Technology Infrastructure

The Language Archive, MPI for Psycholinguistics, Nijmegen, The Netherlands

Arbil: Arbil 2.6, for editing and managing CMDI metadata
Arbil 2.6, for editing and managing CMDI metadata

This manual was last updated in November 2015

The latest version can be found at: http://http://tla.mpi.nl/tools/tla-tools/arbil/

Anastasios Vogiatzis (June 2012)

Claudia Zabeo (December 2012)

Janna Voelpel (September 2013)

Francesca Bechis (July 2014)

Steffen Kaiser (November 2015)

Table of Contents

Chapter 1. Introduction

Arbil is an application for arranging research material and associated metadata into a format appropriate for archiving. Arbil is designed so that it can be used off line in remote locations. The data can be entered at any stage in part or as a whole.

1.1. General Information

This version of Arbil manual focuses on creating collections of CMDI instances. These can be edited and converted into an archivable format to be exported on your local disk or to be submitted to a metadata repository. CMDI instances mainly consist of profiles and components.

Besides creating collections of CMDI instances, Arbil allows you to view all the collections in the archive, or any other remote location (provided that you have access rights), as well as to edit collections from a remote location.

In short, Arbil allows you to:

1) View remote locations;

2) Import metadata from the archive and edit them;

3) Create new collections, with new metadata and media files;

4) Export the CMDI files, ready to be archived.

Changes you make to remote metadata will **<u>not</u>** affect them directly, as Arbil merely allows you to make changes on a **local level**, and not on a remote level.

Arbil is the next generation of the metadata editor, with more powerful user interface and functionality. This manual contains detailed information on how to use Arbil in order to create new collections of CMDI instances and modify existing ones. It does not explain the internal structure of CMDI (for more information on the CMDI metadata, please see the Clarin website [http://www.clarin.eu/cmdi]).

For a less detailed program description, see the Arbil User Guide [http://tla.mpi.nl/tools/tla-tools/arbil].

You can access all the information present in this manual from the application itself, using the **Help Menu**.

We encourage you to provide feedback, comments, and suggestions so that we can improve the documentation to better meet your needs. Also, in case of problems that are not addressed anywhere in this guide, please use the Arbil Forum [http://tla.mpi.nl/forums/software/arbil/].

1.2. Overview

The way you start working on your data is two-fold:

a) either you create a new collection of CMDI instances

b) or you make a copy of an already existing collection from a server and edit it.

Either way, the files you create/edit will be saved on your computer (in the so called **Local Corpus**). This enables you to work off-line in remote locations while still in the field. Changes you make in Arbil will <u>not</u> affect the remote metadata.

Arbil consists of 6 working areas as shown in the figure below:

1) **The Menu Bar** allows you to access the main commands;

2) The **Remote Corpus** contains locations from the Language Archive. This is where you can view remote metadata and choose to import;

3) The **Local Corpus** contains the user's main working area. This is where collections from the Remote Corpus can be imported and edited. It is also used for creating new collections;

4) The **Files /Favourites (Working Directories)**. This is where you can link directories from your computer to Arbil in order to import media files and create favourite metadata items;

5) The **Preview Area** shows metadata of the selected node (in the figure below, it displays the details for "New Corpus"). It changes every time a new node is selected;

6) The **Table Area** is where several nodes can be opened, compared, and modified at the same time (in the figure below, the four nodes from the Local Corpus are placed in a table in this area).

> **Note**
>
> You can customise the size of these panels. To make a window larger or smaller place the pointer on the border and drag it in the direction in which you want the window to expand or contract. It is also possible to hide the preview area by unchecking **Show selection preview** from the **Options** menu.
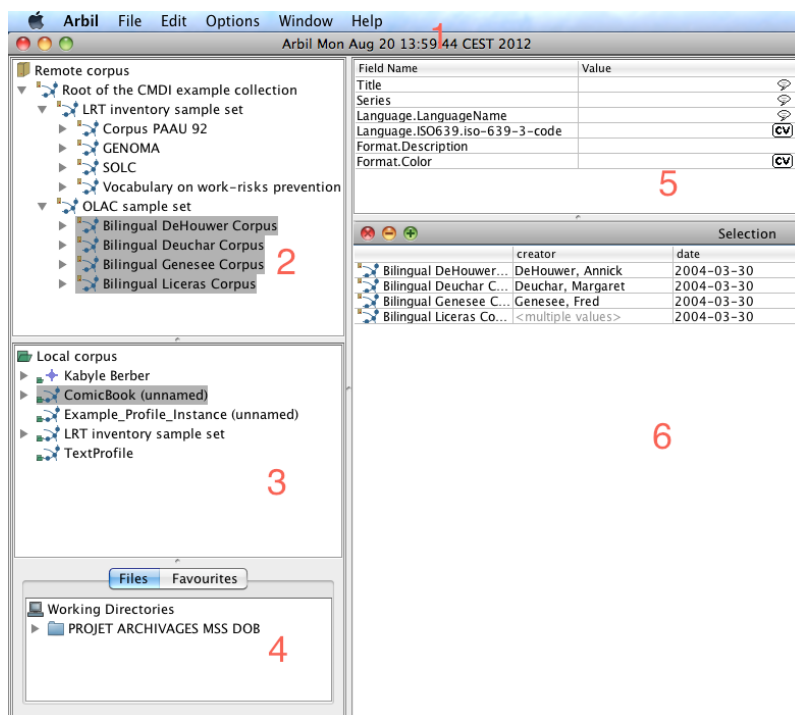
Figure 1.1. Arbil workspace

1.3. Work flow

When working with Arbil, you have to decide whether you are going to work on an existing collection of instances from the Remote Corpus or create a new collection of instances. Assuming that you intend to edit metadata from a remote location and add additional media files, you will have to follow steps 1-6 below.

The Arbil work flow for editing a collection of CMDI instances functions as follows:

1) Browse the Remote Corpus and locate the metadata which you intend to edit; drag the metadata from the Remote Corpus panel into the Local Corpus panel so as to import them.

2) Now that the remote metadata are located in the Local Corpus, you may edit them freely. To edit the metadata, drag the node into the Table Area (see figure 1.1 above, tab number 6) and edit the fields accordingly. This can be done with several nodes at once.

3) In order to add resources to profiles (e.g. media files and annotations), you must set up a working directory which links a folder located on your computer to Arbil. This means that any media files in this folder will appear in the Working Directory area so that Resource files can be added to the nodes in your Local Corpus. To add files to your new collection, drag and drop them from the **Files** tab to the **Local Corpus**. Now you can edit the metadata following step 2.

4) Stored profile components, such as common actors, can be added to the local corpus from the Favourites (either via drag and drop, or via the option **Add from favorites** that you find when right clicking the node in the Local Corpus where you want the favorites to be added.

5) Once your metadata have been edited correctly and your desired media files have been added, you can export the metadata files to a directory (folder) on your computer, ready to be uploaded to a metadata repository.

6) Finally, you can decide to make your metadata public. The procedure for this depends on the institution.
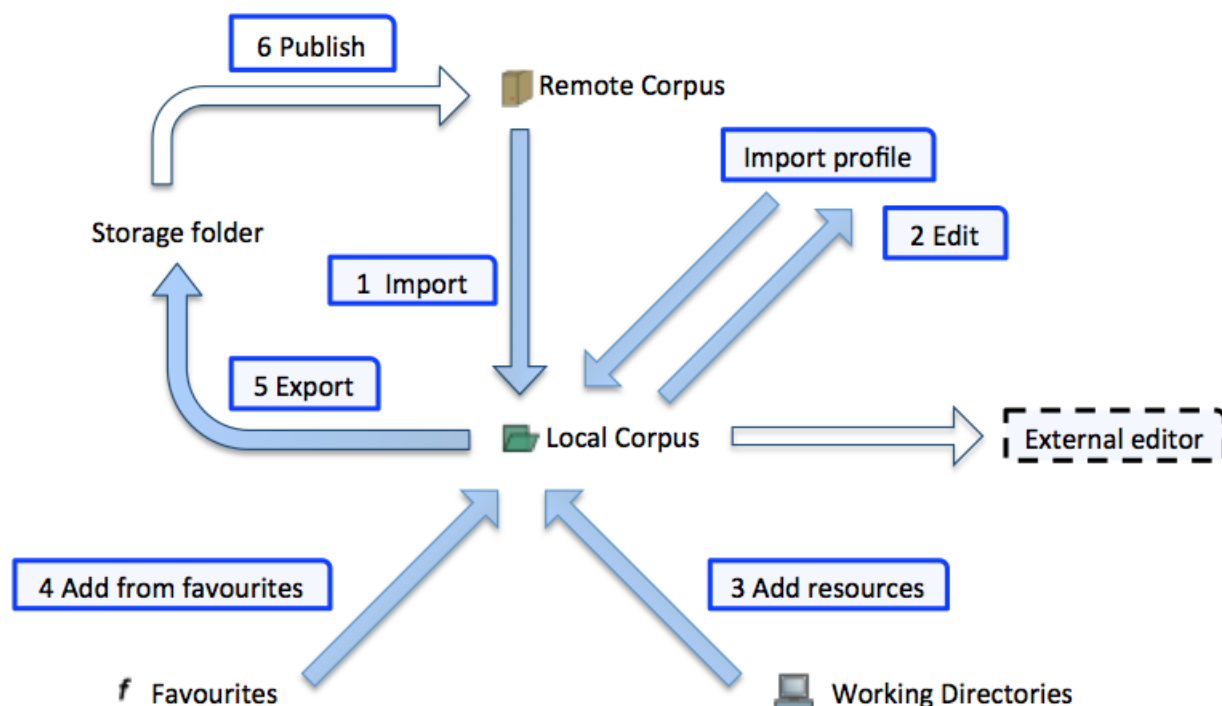

Figure 1.2. Arbil Work flow

In Arbil all work on metadata takes place in the Local Corpus tree. The various operations displayed in the figure above can usually be triggered by right-clicking on elements in the Local Corpus, or via "drag and drop" (e.g. clicking on an element in the Remote Corpus, holding the mouse key, and dragging the element to the Local Corpus, or to a node within the Local Corpus). Media files (i. e. audio and video recordings) and written resources can also be added to nodes via "drag and drop" from the Working Directories tree in the Files tab. Only for importing metadata files from the file system (e.g. previously exported metadata) to the Local Corpus root, you have to access the File menu in Arbil's menu bar.

1.4. Arbil icons

The following table defines graphical icons used by Abril:

CMDI metadata

Attribute

The top node of the Remote Corpus tree (read only)

The top node of the Local Corpus tree (contains local files)

Top node of the Working Directories tree (short cuts of files in the Local Corpus)

The file is in the Local Corpus (local file)

The file is on the remote server (read only)

The file is located in the local corpus but has been imported from the remote server

Closed node (sub-components are folded)

Open node (direct sub-components are displayed

node that contains data

Resource without specified resource link

(2) CMDI node with subnodes; number of subnodes indicated in the brackets.

Node has optional fields but contains no data

Written resource

Video resource

Audio resource

Image resource

Resource is locked and can only be opened by users with access rights.

Resource is unlocked for the user

Info file

In resources: archivable file of an unknown type

File that is not an archivable format and therefore cannot be added as a resource

f The node has been added to the list of the favourites and can be duplicated via the favourites add menu

Node is currently being loaded

File not found

Cosed controlled vocabulary

Language selection on field (e.g. Description language)

Chapter 2. Starting Arbil

When you open Arbil for the first time the *wizard* appears and gives you the ability to configure Arbil to your needs. The first window (Figure 2.1) is a welcome window, press *Next* to move to the second step of the wizard. In this window (Figure 2.2) you select the metadata format i.e. IMDI or CMDI; you can choose one of them or both. In this manual we will assume you will select just CMDI. The third window (Figure 2.3) gives you the ability to choose the corpora and/or the profiles that will appear in the **Remote corpus** box. If you know the URL[1] of the remote metadata that you will be working on frequently, you can add it to this list so that it will be available on the root level of the tree in the Remote corpus box. After finishing with **Remote locations** press **Next** (Figure 2.4) and then **Finish**. You can reopen the wizard by choosing **Run setup wizard** in the **Help** menu.

Welcome to Arbil!

This wizard will help you with the initial setup of Arbil based on the way you plan to use Arbil.

Press **Next** to go to the first step of the wizard.

Figure 2.1. Welcome to Arbil

Select a metadata format

Please select the metadata format you plan to use. Both metadata formats can be used simultaneously in Arbil, but most users will only use one of the formats below. Selecting only one option will *not* permanently disable the other in any way.
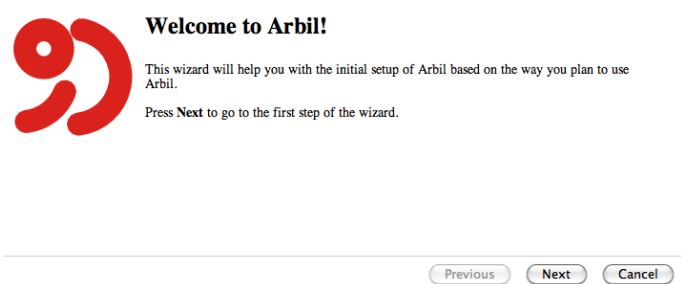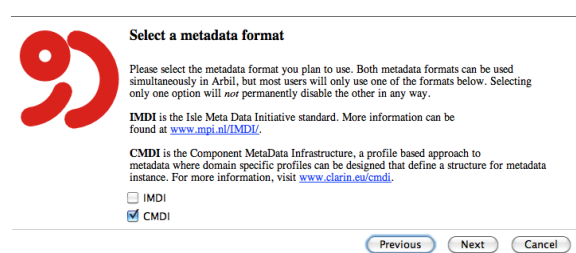
IMDI is the Isle Meta Data Initiative standard. More information can be found at www.mpi.nl/IMDI/.

CMDI is the Component MetaData Infrastructure, a profile based approach to metadata where domain specific profiles can be designed that define a structure for metadata instance. For more information, visit www.clarin.eu/cmdi.

☐ IMDI
☑ CMDI

Figure 2.2. Choosing CMDI

Select your CMDI profiles

Please select the CMDI profiles that you wish to use.
You can always change your selection by selecting 'Templates/Profiles' from the Options menu.

Select the profiles you would like to start using from the list below. If this is the first time you run this Wizard, the profiles will have to be download first. In this case, it will take some time before the list appears.

If you hover over any one of the profile names, you will see a description of that profile.

Clarin Profiles
[Add URL] [Add File]
☐ cmdi-virtual-collection
☐ collection
☐ collection
☑ ComicBook
☐ component-dc-terms
☐ component-dc-terms-modular
☐ CorpusProfile
☐ DBNL

[Reload Clarin Profiles]
☑ Only load profiles selected for manual editing
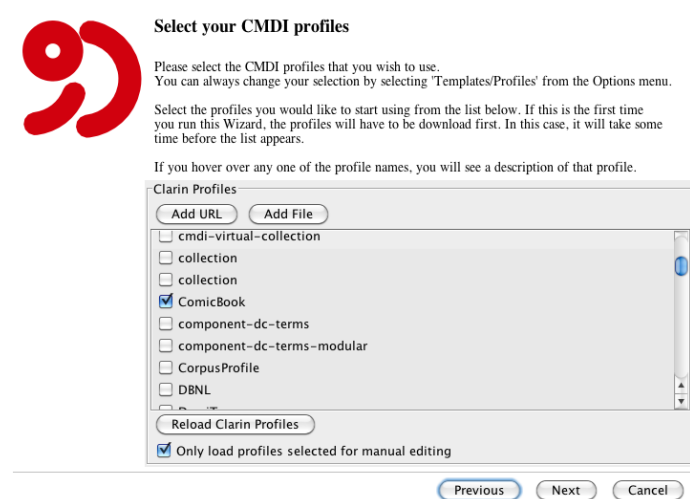
Figure 2.3. Select CMDI profile

---

[1] You can copy the URL of a corpus by selecting the node in the remote tree in Arbil when the wizard is not showing, right click and select 'Copy'; or get it from the IMDI browser by navigating to the corpus, right click the node and select 'Create bookmark' and copy the URL link. Then you can paste it into the box in the wizard.
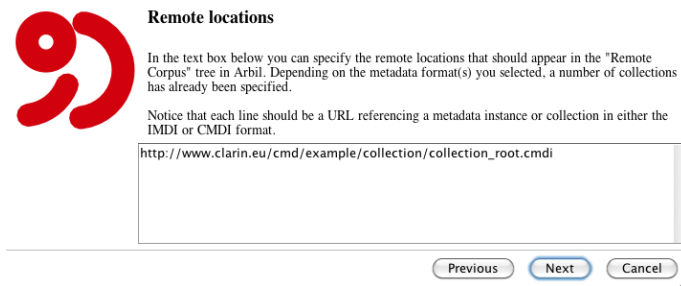
Figure 2.4. Remote locations

Arbil can be started in two ways, either via the webstart or by downloading the installer (see here [http://tla.mpi.nl/tools/tla-tools/arbil]). Note that the webstart does not work for Ubuntu; in this case you will have to download the installer version.

2.1. Language selection

Arbil 2.5 allows you to change the language in which the application will open and work. At the moment, you can choose among: English (either the United Kingdom or the United States variants), German, Italian and Spanish.

When you start the application for the first time, a dialog window will come out asking you to choose a language. After having chosen and selected OK, Arbil wizard will open in the language you have selected (see figures below).

However, you can always change the language via the option **Select language** in the Help menu, which will open the same dialog window as before.

Figure 2.5. Language selection

Figure 2.6. Arbil wizard

Chapter 3. Set-up

Note

> Remember that if, once you have set up your Arbil page, you realize that the font size is either too small or too big, you can resize it as you like it via the options included in the **View** menu, i.e.: **Larger font**; **Smaller font**; **Reset font size**.

3.1. Remote Corpus

The Remote Corpus contains all the remote metadata. Currently it is not possible to add anything to remote corpora or edit it in general using Arbil. It is primarily used to view data and also allows you to import metadata into the Local corpus (discussed in Section 3.2) where they can then be modified.

It has a tree-like structure in which each node contains multiple sub-branches. To open any of them, click on the bullet to the left of the node representing the branch or double click the node itself. This can be repeated for the sub-branches as well. In this way you can view the whole structure of the corpus. If you click the bullet again, the node will fold back.

3.1.1. Tree structure view

As far as CMDI documents are concerned, it is possible to switch to a verbatim representation of the latter in the tree. This means that the structure of CMDI documents will be shown exactly as it is, with no kind of visual simplification. In order to have such a representation, select **Show verbatim XML structure in tree for CMDI** from the **Options** menu.

3.1.2. Adding and removing Remote Locations

If you want to use remote CMDI metadata that are not available under remote corpus, or want to link them at root level, you can add them yourself, provided that you know the URL address of this remote metadata. Right click on the Remote Corpus top node and choose **Add Remote Location**. In the pop-up window enter the URL and press OK. The corpus should now appear in the main list of the Remote Corpus window.
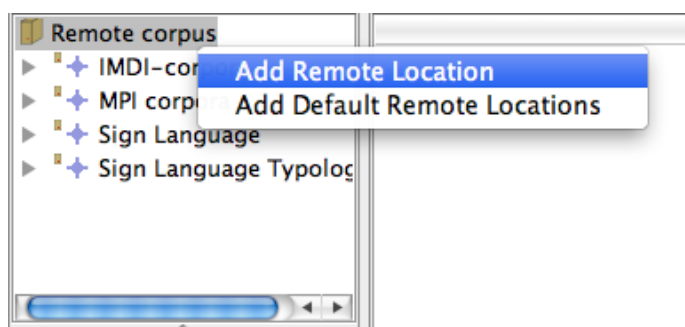

Figure 3.1. Adding a remote location

Since you might not want to use all of the collections of CMDI instances in the Remote Corpus, it is also possible to remove any of the locations of the Remote Corpus at the root level so as to only display the collections that you need. This can be done at the root level only (see figure 3.1) by right clicking on the node (e.g. for MPI corpora) and selecting **Remove Remote Location**. If you have removed the default location, it is always possible to re-add them to the Remote Corpus by right clicking on it and selecting **Add Default Remote Locations**.

To add a subnode to the root level, browse to the desired sub collection, right click on it and choose **Copy**. Then go back to the top node of the Remote Corpus and paste the URL in the dialogue box. The sub collection will now appear in the Remote Corpus tree.

The corpus data on the Remote Corpus may sometimes be updated by the researchers. To be sure that you are working with the newest version of the available corpus, you can update it manually on the nodes that you believe have been altered. To do this, right click on the specific collection you want to update and select **Reload**.

3.1.3. Add to tree root

It may happen that the branch you want to work on is located at a very low hierarchical position in your remote corpus. To avoid clicking on every single node in order to get to the one you need, you can select this one, right click on it, and select the option **Add to Tree Root**. A copy of it will be created and placed at the top of the Remote Corpus tree, so that you will not have to look for it anymore (see Bilingual Liceras Corpus example in the figure below).
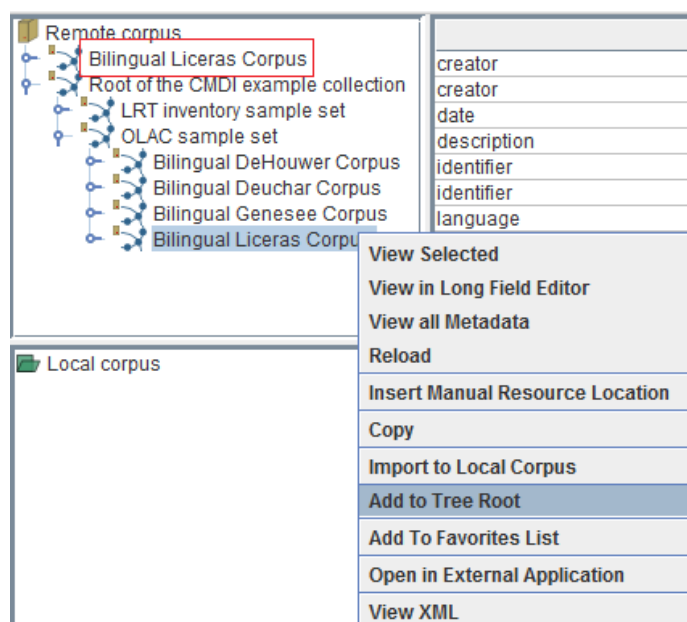

Figure 3.2. Add to tree root

3.1.4. Viewing Remote Corpus Information

In general, the easiest way to view the information of a node is to drag and drop it from the remote corpus into the grey working area on the bottom right. It is also possible to right click on the selected node and choose **View Selected**. In either case, a window pops up with the desired information listed in tables. This also enables you to view multiple nodes at the same time.

To view *multiple* nodes, simply hold the Ctrl key on the keyboard while clicking on the desired nodes. Then drag them to the working area (or right click and select **View selected**). The figure below shows it.

> **Note**
>
> If you have an existing open table, it is also possible to add additional tree nodes of any form (i.e. collections and files). Simply drag any unselected node from the tree structure and drop it on the table. It will be automatically added to the table as a row.
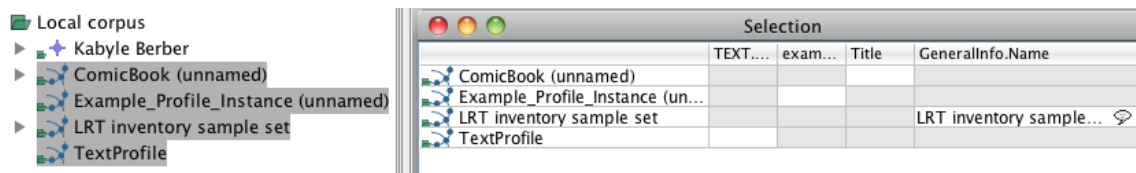
Figure 3.3. Viewing multiple selected corpora

There are also two other quick ways for viewing corpus information:

(1) either you let your cursor rest above a collection or a node without clicking. Information about the collection will be displayed in the tool tip under the cursor;

(2) or you click on a corpus node. Information, if available, will be displayed in the top right frame.

When viewing information on multiple nodes, each row corresponds to one node while each column corresponds to one metadata field. By clicking on any of the column names, you can easily sort rows alphabetically within a given table.

In a table that includes several different types of nodes notice that not all the elements are described in terms of the same field; for example, there are data categories that belong to one collection but not to another. Such data fields are marked in grey and are not available; the same applies to the Local Corpus (see Section 3.2).

Figure 3.4. Not editable cells in a table

In some cases you will see that, instead of a value, in a cell in the table there is a caption: <multiple values>. This means that there is more than one value for this data category for this particular node. To see these values click on <multiple values> and a new window will pop up with the information shown.

3.2. Local Corpus

The Local Corpus ⬚ is located beneath the Remote Corpus panel. Remember that the Local Corpus is the main working area and that changes made in the Local Corpus will not affect the remote matadata. The Local Corpus serves two main functions: 1) editing existing collection, 2) creating new collection. Section 3.2.1 explains how Local Corpora can be modified by (re)editing the data from the Remote Corpus. Section 3.2.2 explains how to create and subsequently modify new collections.

3.2.1. Import to Local Corpus

If you have decided to work on a corpus from the archive you first need to import it from the Remote Corpus to your Local Corpus, and then you can edit it. To do so, select the node/session from the Remote Corpus and simply drag the item and drop it into the local corpus icon ⬚. You can also right click on the corpus node you want to import and choose **Import to Local Corpus**.

In the dialogue box that pops up you may want to check the option **Import Resource Files (if available)**, through which, as the name itself says, both the node and the possible resources that such node contains

will be imported into the Local Corpus. Please be aware that: 1) Arbil cannot deal with resources that need authentication; 2) copying the resource files may take some time and will use disk space on your computer.

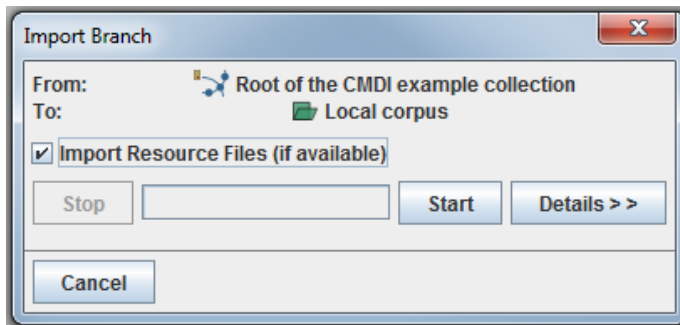Select **Start** to initiate the import process. Select **Stop** if you want to interrupt such process.

Figure 3.5. Show details and options

Once the import process has started a message can appear, asking you whether you want to overwrite an existing local corpus/session node from the same original location. When is this option useful? Consider the likely scenario that you have imported a node from the Remote corpus and started editing it. If you want to reset the changes, you can reimport it from the Remote Corpus and answer "Yes" to the question "Overwrite?". You can apply this option to every branch of the corpora.

You may want to check the option **Remember my choice**, so that the message will not come out again for the other already existing files that Arbil will encounter during the import process.
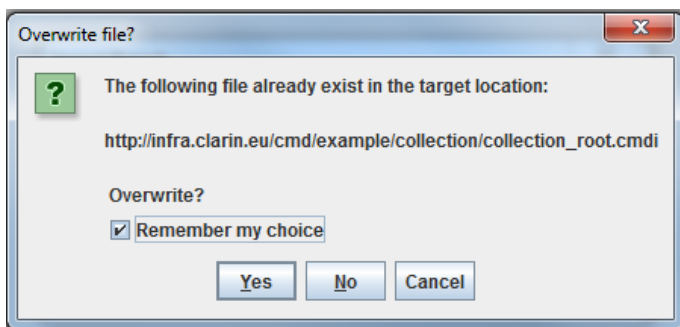

Figure 3.6. Overwrite message

Once the import has finished you will see a window telling you how many Metadata Files have been processed. If you select **Details**, you will see all the possible errors that have occurred during the process.
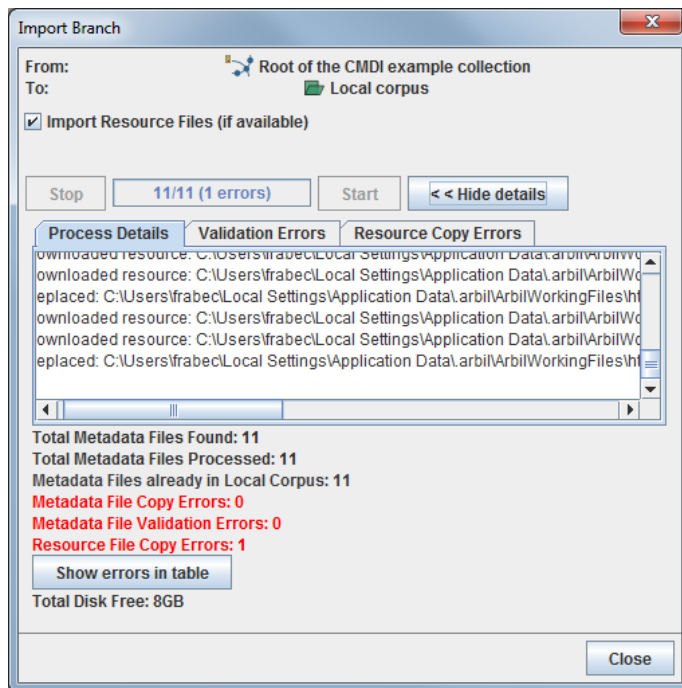
Figure 3.7. Import completed

3.2.2. Creating profile instances

Instead of copying an existing collection or corpus nodes from the Remote Corpus, you can also create your own metadata instances. Moreover, you can add an existing profile; with multiple profile instances it is possible to build a tree structure.

To do so, select the top node of the Local Corpus (corpus node), right click it and select the option **Add** and then you can choose the profile or component(s) you want to add.

3.2.3. Creating (and describing) hierarchical collections

In Arbil CMDI you can also create hierarchical collections of instances. The easiest way to do so is to create the parent and the child nodes **separately** in the Local Corpus, as if they were two different examples of instances. Once both are there in the Local Corpus tree, simply drag the child node into the parent node. A message will appear asking you whether you want to move the node X to the node Y. After selecting **Move** you will see the Local Corpus tree structure changed (see figures below).
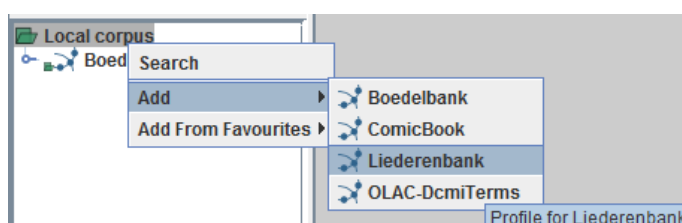

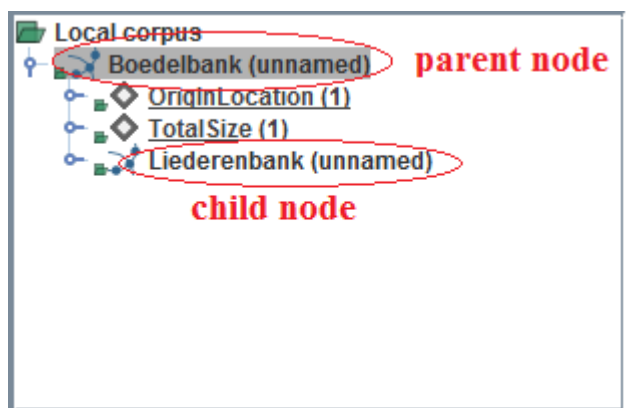**Figure 3.8. Creating parent and child node**

Figure 3.9. Creating a hierarchy

> ![Note icon] **Note**
>
> The recommended Clarin profile to use in order to describe collections of instances is clarin.eu:cr1:p_1284723009187 [http://catalog.clarin.eu/ds/ComponentRegistry?item=clarin.eu:cr1:p_1284723009187].

3.2.4. Local Corpus Storage Directory

When Arbil is started for the first time a .arbil directory should be automatically created on your local machine. In it, and more specifically into a subfolder named ArbilWorkingFiles, the content of your Local Corpus should be located and saved. However, the location of your Local Corpus can always be changed. To do so, go to **Options** > **Local Corpus Storage Directory** > **Move Local Corpus Storage Directory...**.

> ![Note icon] **Note**
>
> It may happen that Arbil detects multiple storage directories on your machine (which may be due to an older version of the application conflicting with a new one). In those cases you will be notified through a warning message, asking you what you are going to do with such storage directories.

3.3. Working Directories

Beneath the Local Corpus window, you will find two different tabs: **Files** and **Favourites**. This section describes the Files tab and the functioning of its **Working Directories**, while Section 3.4 below focuses on the Favourites tab.

The Files tab is responsible for linking a folder from your computer to Arbil so that media files can be transferred into your corpus. It is advisable to have one or more folders on your computer that contains all the necessary files that you would like to upload to your corpus. This way, you can simply link this folder and retrieve all the necessary files.

The option Files is selected by default. Its Working Directories are meant to keep at hand the resources that you want to add to your corpus (such as annotations, audio files, pictures etc.). To add a new Working Directory, right click on the Working Directory node and select **Add Working Directory**. Now browse to the directory you want to add, select it and click again on **Add Working Directory**. In the figure 3.8, the directory "Recursion pics" is about to be added to the Working directories, after which you will be able to access the media files present in this folder.
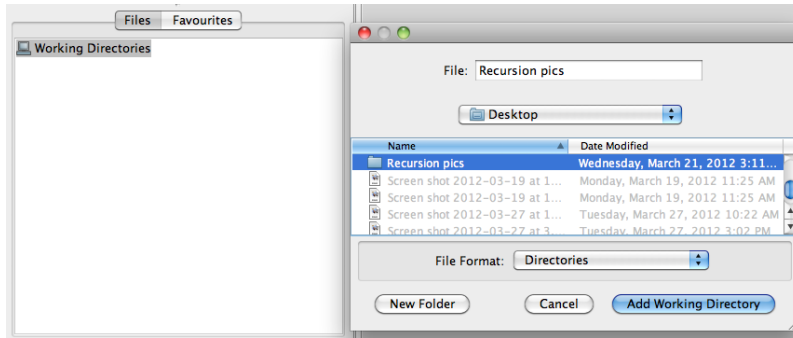
Figure 3.10. Add working directories

After linking a directory to Arbil, you will have access to the folder. To import the media files (e.g. Droste.jpg) to a session of a corpus, simply drag and drop them into the desired session.
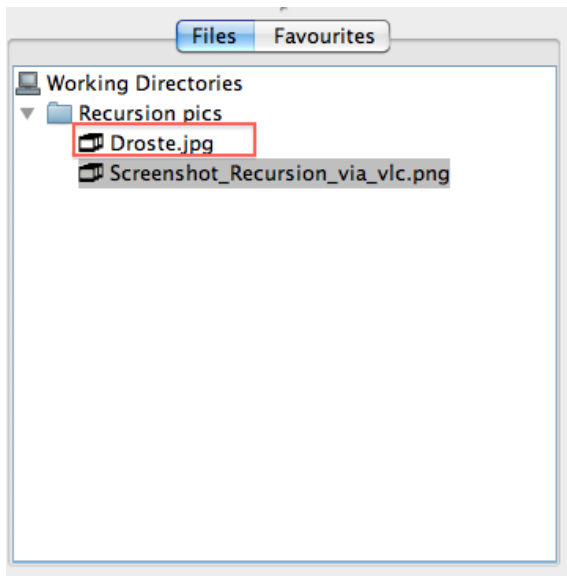

Figure 3.11. Linked folder

> ### ⚠ Note
>
> If you right click on Working Directories, you will also see that it is possible for Arbil to show you all hidden files (option **Show hidden files**). This option is useful when you want to include collection files which are normally hidden in your file manager.

Browsing through the Working Directories is very similar to browsing through the Remote or the Local Corpus. To open/close a folder click on the bullet on the left of its name or double click on the name itself. If you have added or modified the content of a directory outside Arbil, you can update such content by right clicking Working Directories and selecting **Reload**.

If you want to view the file information you can follow the instructions described in Section 3.2.2. (Viewing Remote Corpus Information). Remember that if you place the cursor over a file, an information window will appear that will give you a file description, including the information about whether the file is archivable or not. Archivable files can be uploaded to the server. Image 3.7 shows an example of an archivable file:
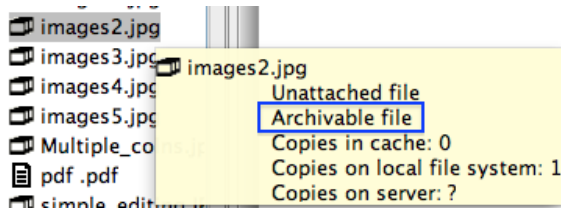
Figure 3.12. Archivable file

Editable files from the Working Directory can also be opened by external applications. Right click a given file (e.g.. audio file, picture) and select **Open in External Application**.

If the file is not archivable, the   icon will appear. If you hold the mouse pointer on the file, you will see the following information:
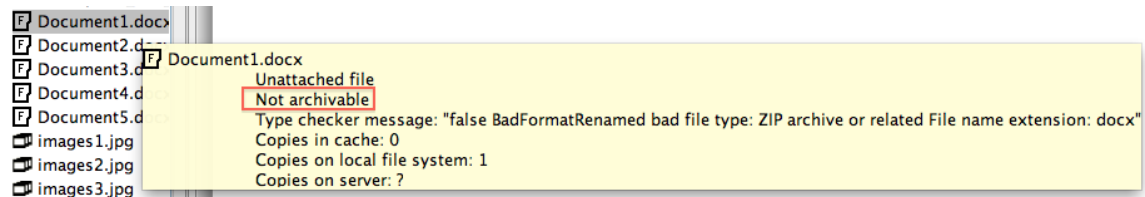

Figure 3.13. Unarchivable file

For the list of accepted file types and formats, check the LAMUS manual [http://www.mpi.nl/corpus/html/lamus/apa.html].

A file that is not supported by Arbil can be kept in Arbil anyway and still be added to a Session in the Local Corpus, thus overriding the decision of the internal type checker. To do so, right click on the unarchivable file and select **Override Type Checker Decision**; a pop up window will appear asking you to say whether the file is a **MediaFile** or a **WrittenResource**; the   will then change into an icon of either a media file or a written resource file, and you will be able to add it as a resource file into the Local Corpus. To open the file right click it again and select **Open in External Application**. Beware that overriding the type checker decision can cause your corpus not to be accepted into the archive (rejected by Lamus).

In addition, if you have modified your uploaded file externally, you may want to check whether it is still accepted by Arbil. To do so, right click the file in Working Directories and select **Force Type Checking**. If something has happened to your file you will see its icon change.

3.4. Favourites

The tab **Favourites** is meant for storing elements that you would like to re-use in new node, such as *actors, language elements, locations*, etc. Node elements that can be saved in Favourites are CMDI instances, or any subnodes. When taken to the favorites tab, such elements are placed in their corresponding subcategories; i.e. if you drag an author, this will be automatically located under the category "Author". If you have a number of elements in your Favourite List and you are not certain what the content of a given element is, you can always place the cursor over it and wait for the tool tip (information window) to appear.
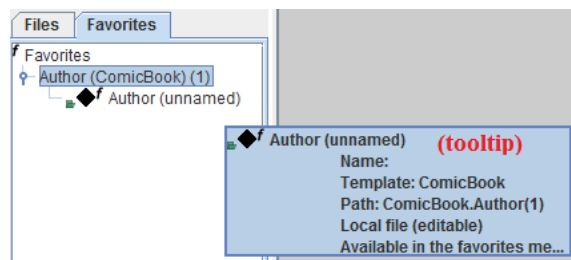

Figure 3.14. Favourites List

In order to add an element from the local corpus to the Favourites List, right click its node and choose **Add to Favourites List**. The element now appears in the Favourites List. You can also just drag them from the corpus tree structure, right down to the Favourites List. If this element is needed again, you can place it into a preferable location in your Local Corpus either by dragging it from the Favourites list or by copying and pasting it in one of the Local Corpus (sub)branches.
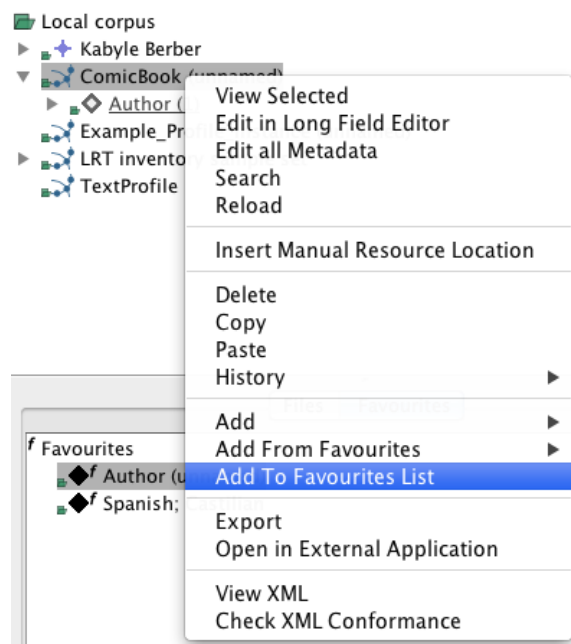

Figure 3.15. Adding to favourites via menu

Once you have completed your Favourites list, you can start using it. Let's say you want to add a particular Author to a session in your Local Corpus: in order to do so choose the session in question, right click it, select **Add from Favourites** and click on the desired author. Otherwise, you can simply drag and drop the author from the Favorites tab to session in the Local Corpus.
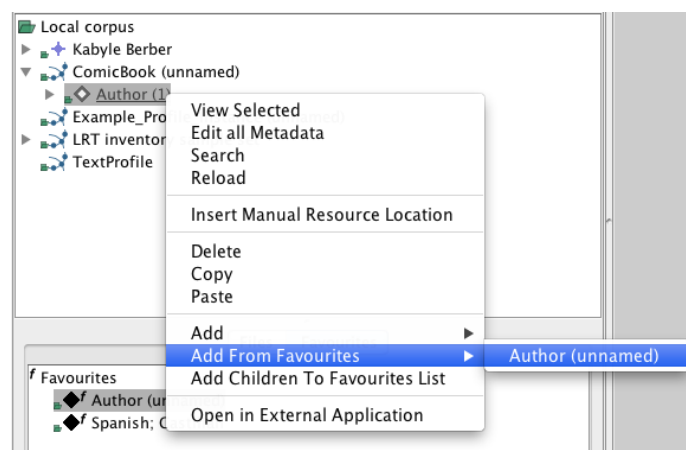

Figure 3.16. Adding from Favourites to a component

Note that the items in the Favourite List can be viewed and edited in the same way as the elements of the Local Corpus (i.e. via the options of the right click context menu). This means that if you make changes to your favourites you will have to right click on them and select **Save Changes to Disk** to make them permanent (or use Ctrl/cmd+S). The favourites which have been edited turn blue, which indicates that they contain pending changes to be saved locally.

Favorites can also be exported and/or (re)imported. In order to **export** your favorites: right click on **Favorites** (in the Favorites tab) and select **Import/Export Favorites**. A window like the one in the figure below will appear. Select the favorites you want to export, then **Export selection**, and after having chosen the location where to export select **Save**. As soon as the export process is complete you will see a message confirming it and the folder containing the exported files will open.
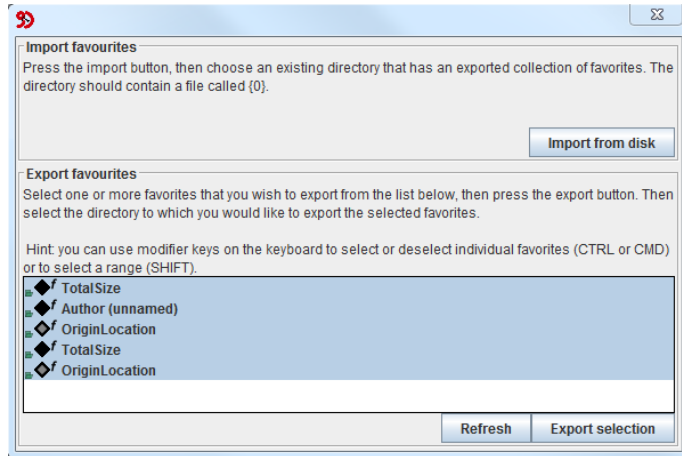

Figure 3.17. Export/Import Favourites

As you can see from the image above, if you want to **import** some favorites into Arbil you have to follow the same instructions as the ones explained for the export: right clich on favorites > select Import/Export Favorites. This time, though, choose **Import from dick**. Choose the (.config) file that you want to import and select **Open**. You will see a message confirming the successful import, while in the Favorites tab you will see the just imported files.

3.5. Managing Selected Profiles

From the **Options** menu, choose **Templates and Profiles**. A list of the available CMDI profiles will appear. You can select the profiles you need from the list or you can add new ones.
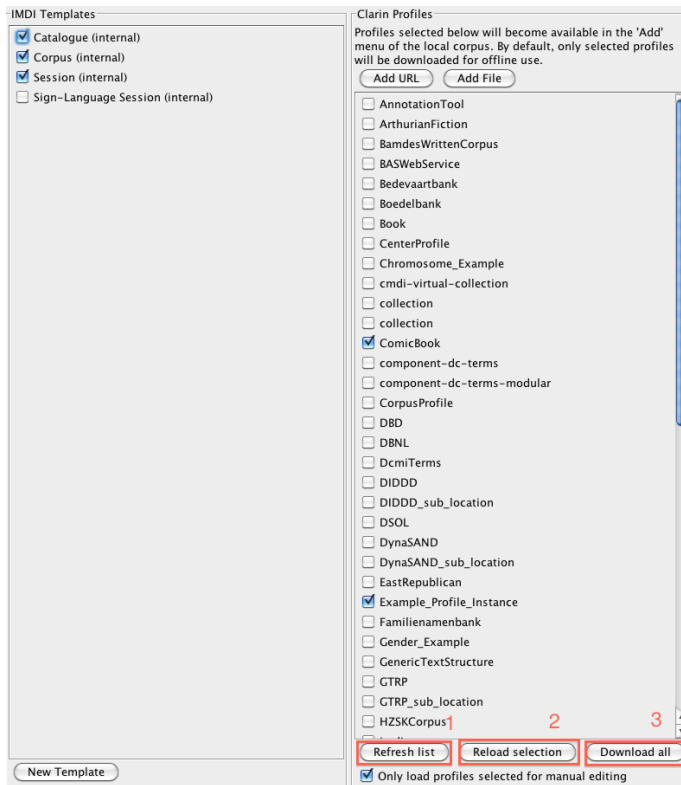
Figure 3.18. Available Profiles

The button **Download all** (see number 3 in the above picture) allows you to download all the available profiles, so that you will be able to see them also without an internet connection.

> ⚠️ **Note**
>
> Another way of having the profiles available also offline is to *select them from the list when online*, since they get downloaded as soon as they are selected.

When you add new profiles manually, or there are changes on them, you need to reload the list through the button **Reload Selection** (number 2 figure above) in order for the changes to be visible. Note that you have to *quit and restart* Arbil to fully validate the reload process.

With the button **Refresh List** (number 1 figure above) the list of profiles available in the Component Registry will check for updates, so that you will get the latest list of profiles form the server. In this case, unlike with the "Reload" commad, you do not have to quit and restart Arbil in order for the changes to be visible.

In order to add a new profile, select the option **Add URL** (it is located above the list of profiles; see figure above) and paste into it the URL of the profile you want to import. You can find the URL, for instance, in the **Clarin** website, under the **Component Registry**. Note that only published profiles appear by their name; otherwise, they are displayed in the list by their ID (something that starts with clarin.eu).
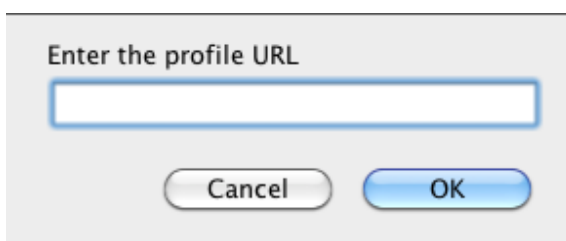

Figure 3.19. Add URL

In order to use a profile, this needs first to be selected through the check box. Then, if you right click on the Local Corpus node and go to the Add option, another menu will open with the names or IDs of the profiles that are selected in the list. Choose the one you need. This will be imported and it will appear in your Local Corpus.

Note

A profiles' components can be edited directly in the Component Registry of the Clarin website. If you do so, remember that you have to add the URL to the Arbil list of profiles again (the process is the same as the one explained above: Options menu > Templates & Profiles > Add URL > Reload > quit and restart). Once the modified profile has been added, if you go to the Local Corpus, right click it, and select Add, you will see two profiles with the same name but with different details/structures (the old one has not been replaced).

Chapter 4. Editing the structure of an instance

This chapter focuses on how to develop the structure of the Local Corpus.

4.1. Adding components and/or elements

Once you have created/imported a collection, you can add new components or elements to any corpus node of the Local Corpus. To do so, right click on a corpus node, scroll down to Add and choose the components or elements that belong to the imported profile(s). Notice that in the figure below the profile "Comic book" has been selected. The components and elements which are displayed belong to that specific profile. Different ones will appear accordingly for each profile you decide to use.

Notice how some menu items have the ◆ icon, whereas others have a ⌐ icon. The ◆ icon represents a component that will be displayed in the hierarchy whereas the ⌐ icon adds a new field (discussed below).
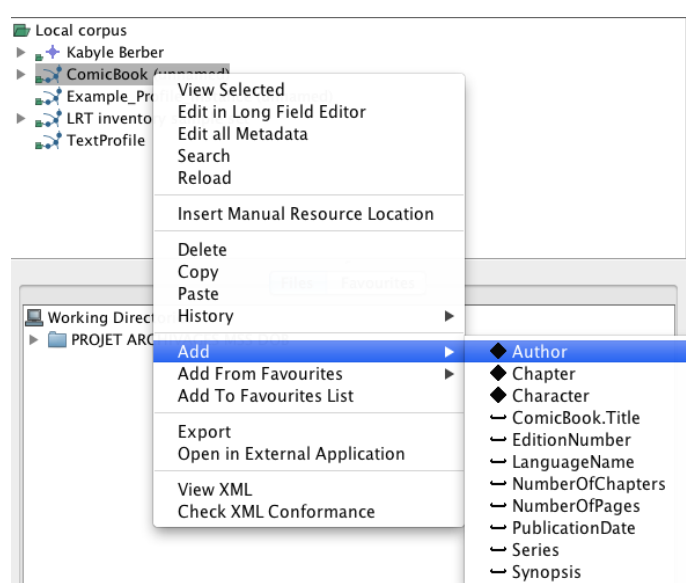

Figure 4.1. Adding a new component or element

4.2. Adding more field elements

Generally, every profile contains a particular set of fields which can be edited. Adding ⌐ elements implies adding a new metadata field to the selected node.

Figure 4.2. Standard field name

Using the example of our ComicBook profile, it consists of field names which include one field per title. You might want to add an additional 'Title' field. To do so, right click on a node, Add > (e.g.) ComicBook Title. Some fields can be added multiple times, others can be added only once.

Figure 4.3. Add new metadata field
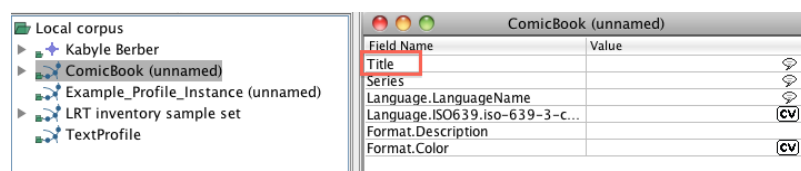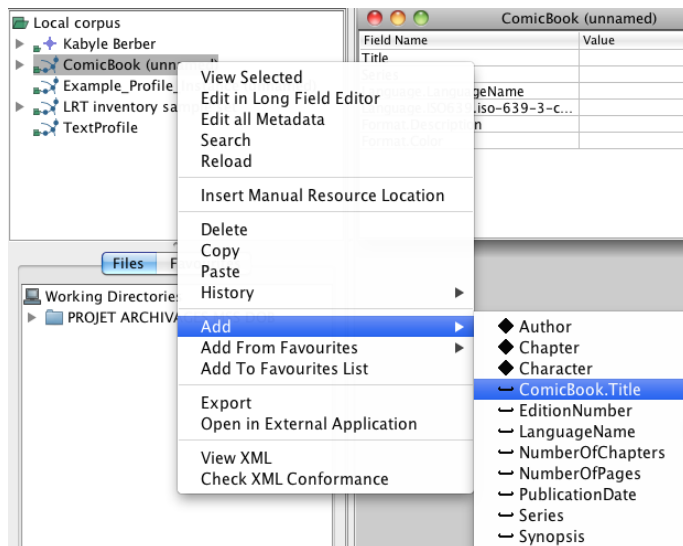
4.3. Adding resource files to a metadata document

The easiest way to add resource files (e.g. audio, video etc.) to a metadata document is to drag and drop a given element from the Working Directories to any existing node in a metadata document. There it will appear as a child node of the selected node.

Note that not all file types are archivable. If a file is not archivable, you will not be able to drag and drop it (any problem with the file will be illustrated by the   icon).

Figure 4.4. Adding a resource by dragging and dropping

It is also possible to manually add resources from a remote location, i.e. the archive or the web. This can be done by right clicking on a node, selecting **Insert Manual Resource Location**, and adding the URL in the window that appears in the table area.

Figure 4.5. Insert manual remote resources

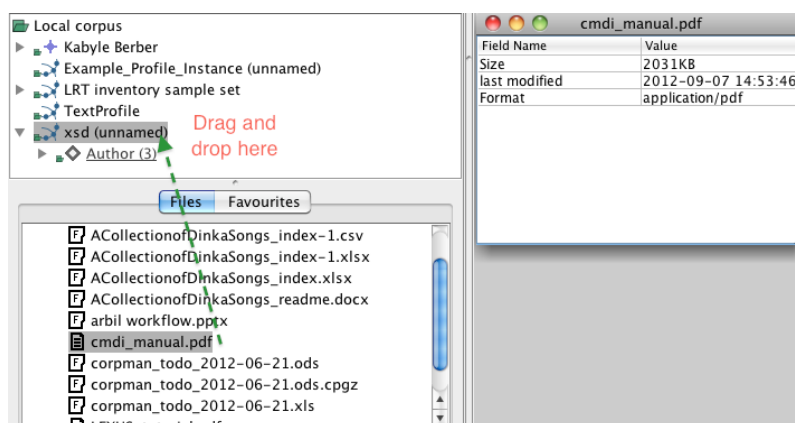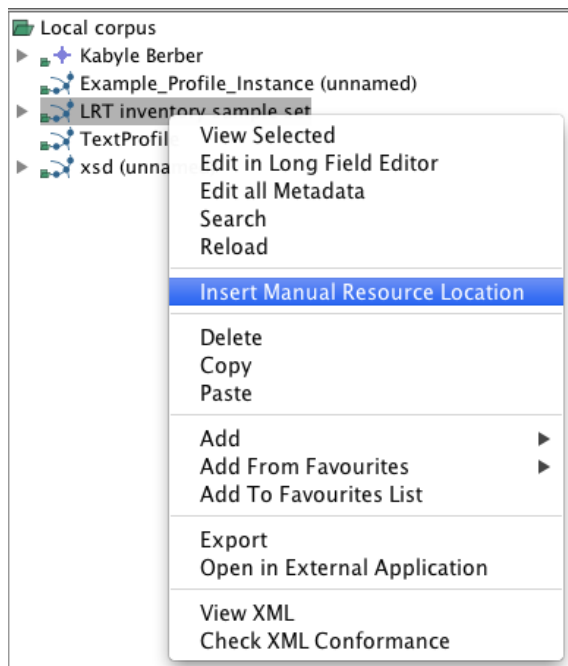When adding multiple resources from the Working Directories to a CMDI file, they will get processed together and appear in a single table in the table area.

4.4. Moving instances

Arbil allows to move metadata from one location to another. To do so, select in your Local Corpus an instance or a collection you would like to move and drag and drop it to the new location. It is also possible to move subnodes of collection, e.g. Actors, across sessions. However not all locations will be accepted, depending on the type of the dragged node.

4.5. Deleting linked instances

In order to delete metadata instances or child nodes, right click on a chosen element and select **Delete**. If you want to remove several elements at a time, choose them all while holding the Ctrl (or Shift) button. Right click on the selected elements and chose Delete. In each case you will have to confirm the deletion.

4.6. Saving changes

The changes to your files are not applied automatically. You have to save them yourself. To save the changes made in your local corpus: 1) right click on the node and choose **Save Changes to Disk**; 2) from **File** menu select **Save Changes**; 3) use the key short cut (i.e. Ctrl+S/cmd+S).

Chapter 5. Editing the metadata

5.1. Editing one cell

To edit a metadata document or a collection just click on it (A), or select it and choose **View Selected** (B). In the window that appears, click on the field you want to change. Type in the new value: typing in an existing cell will start editing for replacing the existing content. Pressing the F2-key on windows/linux or CTRL-U-key on a Mac will start editing without removing content. Press **Enter** to go to the next value. If the value you have typed is too long for the cell, the **Long Field Editor** will open automatically in order to edit this value. (section 5.2)
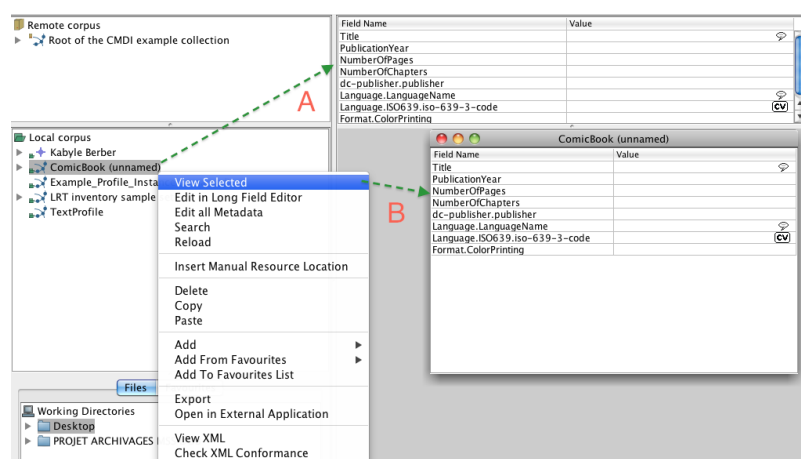

Figure 5.1. Editing

Some of the cells have a closed vocabulary, ⟨CV⟩, (figure 5.2). When you double click the field you need, a drop-down menu will appear showing you available default values from which you can choose the one that fits your data. In controlled vocabulary fields 'type-ahead' functionality is available; to confirm a suggestion, press Enter or type a comma in multi-CV to add another value.
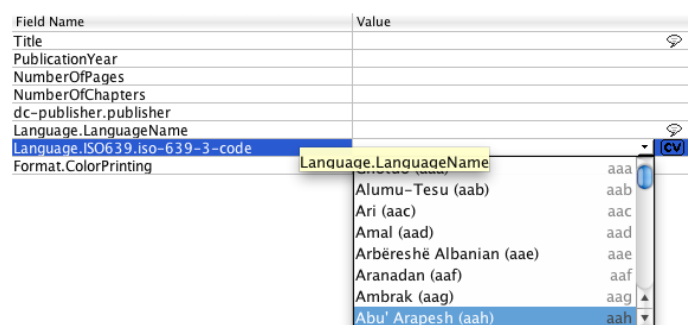
Figure 5.2. Closed vocabularies

5.2. Attributes

It is possible to fill in or modify fields which contain attributes by right clicking on a selected node and choosing View Selected. Once the node has been displayed, the attributes fields are ready to be edited. There is also the possibility of working on them through the Long Field Editor. In case of long descriptions, one or more attribute fields can be edited here.

Note that whether attributes are present or not depends on the profile that has been selected.

Figure 5.3. Attributes

In the tooltip which appears hovering on a field, you can also see information about the attributes of that particular field. The number of potential attributes, as well as the value for the attributes that have one specified are displayed too. For an overview of all the possible attributes, open the field in The Long Field Editor (Section 5.3).
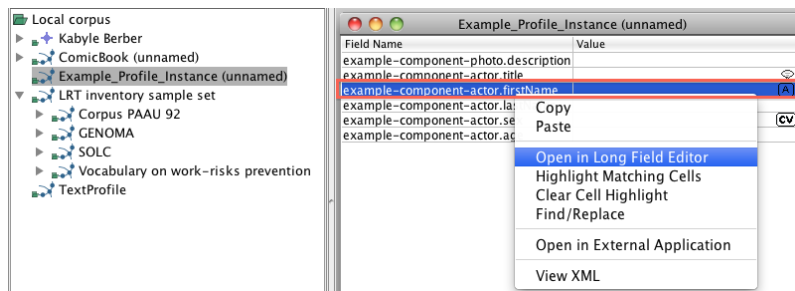
5.3. Long field editor

The Long Field Editor (LFE) allows the user to edit long values, such as descriptions. The LFE works either automatically or manually. If you want to add a long description in a cell you can right click on the field in the table and select **Open in Long Field Editor**. In the window that appears you can add new text and/or edit the already present text. If a field already contains a long value and you wish to edit it, double click on it; the LFE will open automatically.

The LFE can be opened on a selected node by right clicking on it and selecting **Edit in Long Field Editor**. The LFE opens on the first field under the node (generally the field **Name**), you can then navigate in the fields by using the *previous* and *next* buttons (1). The field you are currently in appears on the top of the LFE (2). The text appears and can be edited in the white area (3).

Changes made in the LFE get applied to the field as soon as the LFE is closed or by pressing the *previous* or *next* buttons.
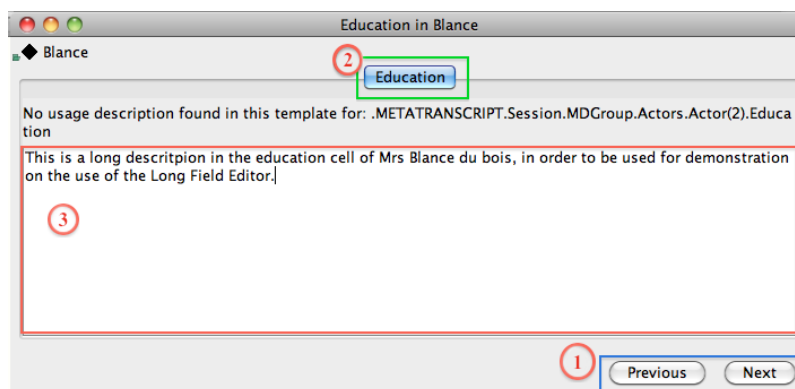
Figure 5.4. Long Field Editor

5.4. Delete field(s)

Fields can also be deleted both for the metadata table of a single instance (see figure below), and from the one containing several nodes. As far as **one single instance** is concerned, you have to open such instance in the table area (either via right click > **View selected**, or simply by dragging and dropping the instance in the main area). Then select the field you want to delete, right click it and select **Delete 1 field(s) "..."**. Note that mandatory fields are NOT deletable; in case of a mandatory field you will not be able to see the delete option.

Figure 5.5. Delete field from single instance

In case of multiple instances in one single table (see figure below), what changes is the layout of the table itself, but the 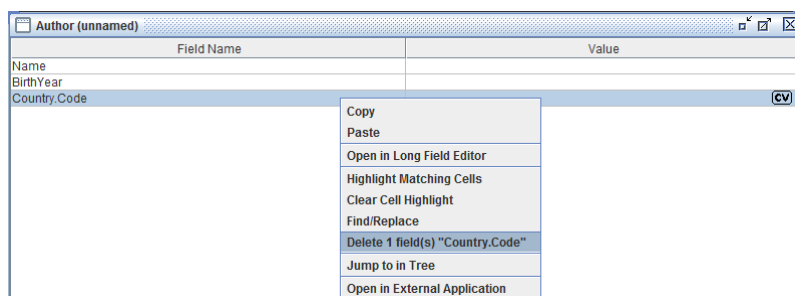deletion process is the same as the one explained above for single instances. Open the instances in the table area, choose the field you want to delete, right click on it and select **Delete field from all nodes**. Note that, as before, mandatory fields CANNOT be deleted. Unlike with single instances, though, here you will always be able to see the delete option, the only difference being that in case of undeletable field a message will pop out telling you that that specific field cannot be deleted from the nodes.
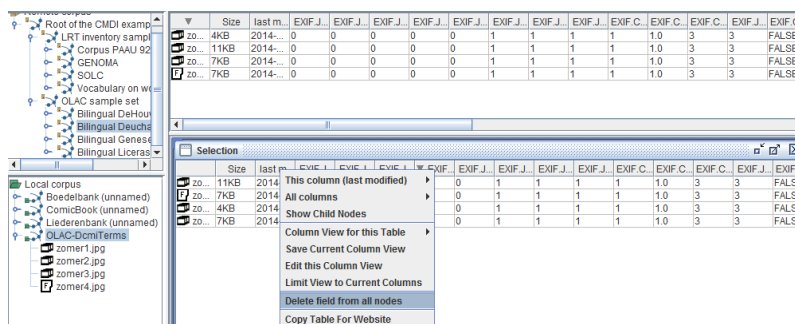

Figure 5.6. Delete field from several instances

5.5. Editing multiple cells

It will often be the case that a piece of information (e.g.. continent, country, address etc.) has to be repeated across the metadata documents. Instead of typing this data in every document, you can copy-paste it at once into all the relevant fields. Choose the relevant metadata documents by holding Ctrl and clicking on them. Then right click one of the selected items and choose View Selected. You will see a window appearing in which you can edit the documents.

To edit the selected metadata documents, take the following steps:

1. Fill one of the fields / choose a default value from the drop-down menu that you need (e.g. Country). If one of these fields is already filled in with the relevant information, skip this step.

2. Right click on one of the selected cell and select **Copy**.

3. Select the fields where you want to paste the copied information (one by one) by holding Ctrl and clicking on them. Right click on one of the selected fields and chose **Paste**.
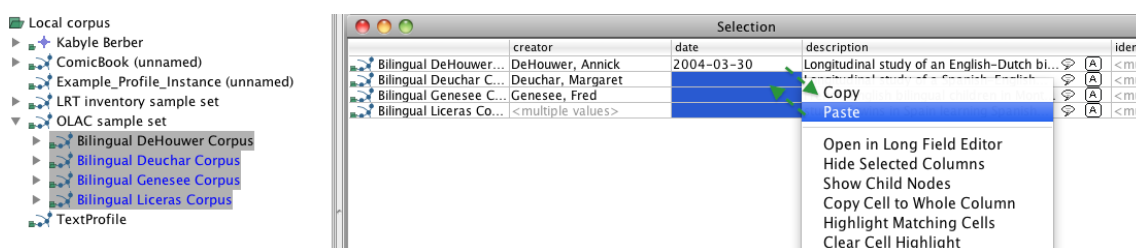

Figure 5.7. Pasting data into multiple cells

If you want to copy a value into **all** the cells of a certain column, select the field you want to copy, right click it and chose **Copy Cell to Whole Column**. All the cells in the column will be changed accordingly. Note, however, that this option is only applicable to one column at a time.
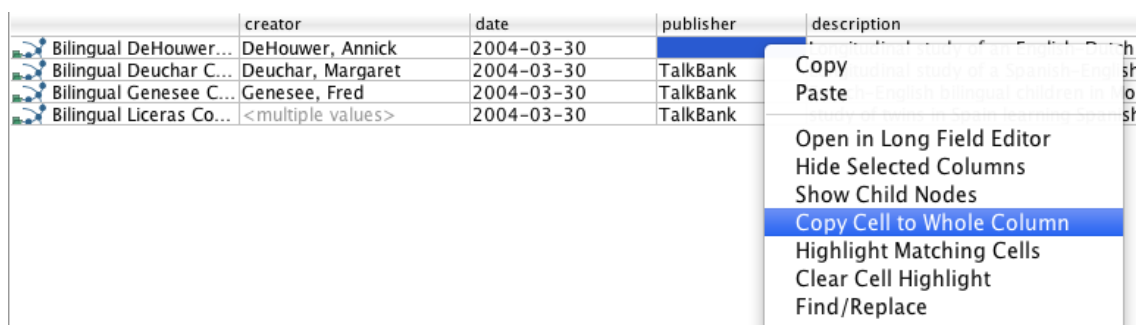

Figure 5.8. Copying the value of a cell to all the cells of a column

It is also possible to copy-paste values to multiple columns simultaneously. To do so, while holding Ctrl (or cmd), select the relevant cells of a metadata file, right click one of them and choose **Copy**.
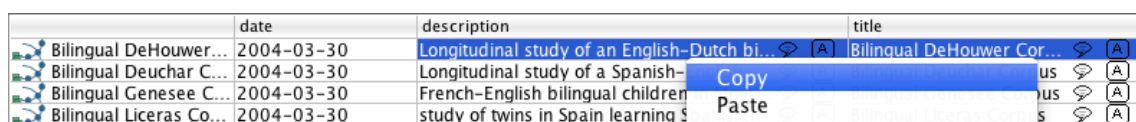
Figure 5.9. Copying values from columns in one metadata file

Now choose the cells within the selected columns into which you want to paste the copied value by Ctrl (or cmd) clicking on them (you can also draw a box over them with the cursor), right click one of the cells and select **Paste**. The values will be pasted into the appropriate columns.

Note that you cannot copy-paste data from different rows, only from different columns.

Figure 5.10. Pasting values from two columns into several metadata files

It is possible to show the entire document containing any specific node that is shown in the horizontal table view. To do so, right click a cell or row and select **Show Context**. This menu item opens the hierarchical 'sub-nodes' view on the root of the document that contains the selected cell or row.

Figure 5.11. Show Context Menu

5.5.1. *Find* and *Replace*

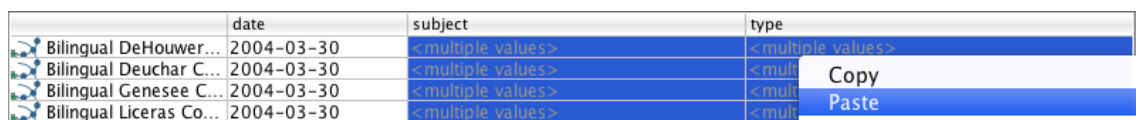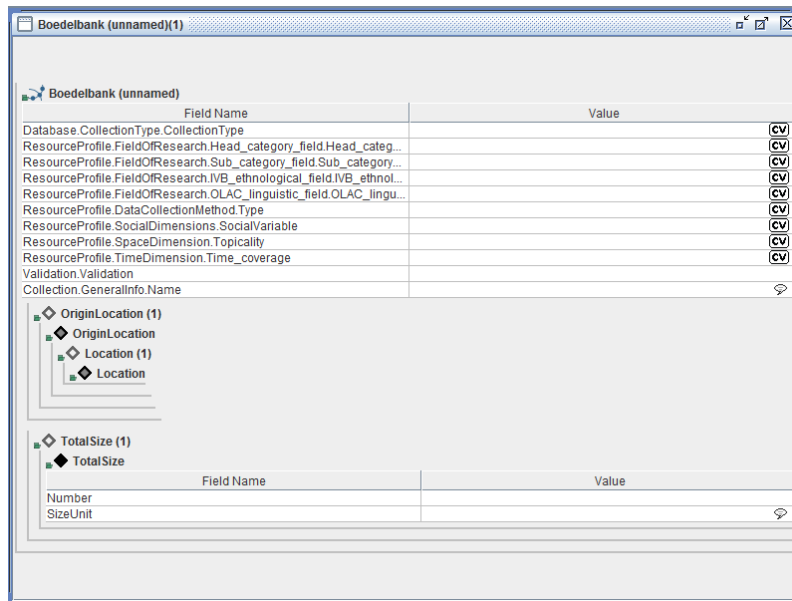Another option in editing cells is the *Find and Replace* option activated by pressing, when you are in a table, CTRL/CMD+F, or by right clicking on a cell and selecting the option **Find/Replace**. In both cases a line appears at the bottom of the table with two boxes (figure below); the left box (marked in red) is used to define the term you intend to find; the right box (marked in green) is used to define the term to be replaced with. By pressing the **Find Next** key the term specified in the first box is highlighted in the table; now you can either replace such term by pressing the **Replace Selected** key, or skip it and move to the next by pressing Find Next. Once you have moved to the next term in the table, you cannot go to a previous term. In that case you will have to manually select a previous cell to start from and continue searching.

Figure 5.12. Find and Replace

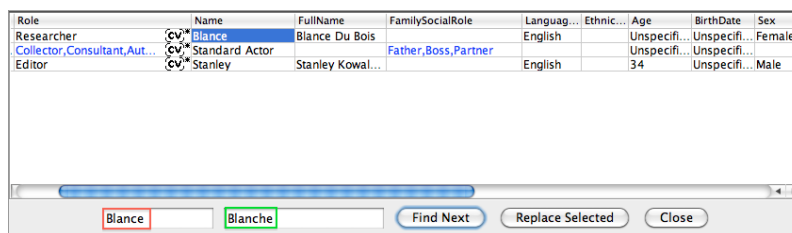## 5.6. Editing All Metadata of a Subnode

It is also possible for the user of Arbil to get access to all metadata under a node in a hierarchical order (figure below), as the order of the (sub)sessions, so as to edit them. To do this, right click on a metadata document on the tree view and select **Edit All Metadata**. Editing can be done in all cells in the way it has already been described.
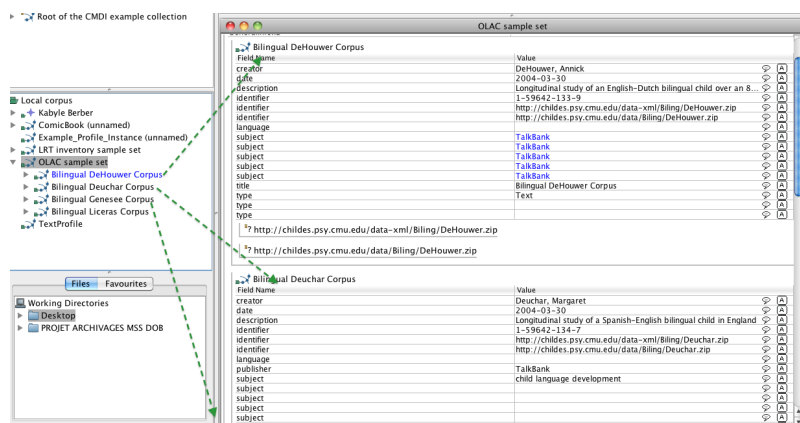
Figure 5.13. Edit/view all metadata

5.7. Customizing Column Views

When viewing/editing multiple elements in a table, sometimes there are many columns to go through. It is possible to reduce them by choosing which columns you want to see or which columns you want to hide. Right click the header of any column, i.e. Name, and select the option **Edit this Column View**. In the window that pops up, you can select the columns that you want to show **Show Only** or hide **Hide** (if you check only the Hide column, the unchecked field names will be shown in the table). It is important to know that the column(s) which are hidden through this particular window, are not removed from the corpus itself.
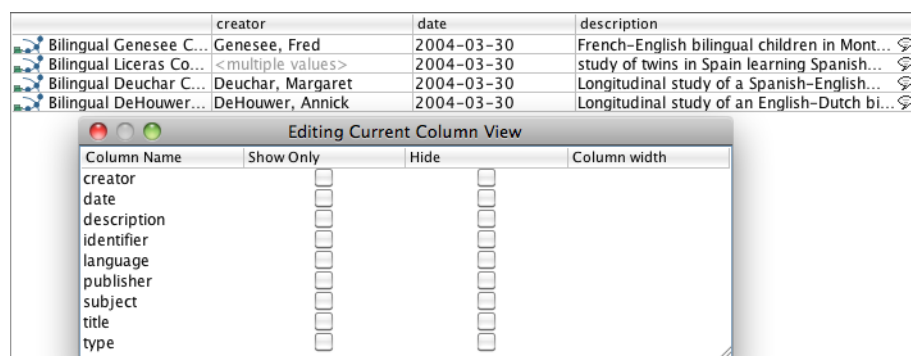

Figure 5.14. Choosing columns

So let's say that you have *Date* and *Description* marked as *Show only* (see figure below). This means that if you would now use this window to check out a node that has other categories than the ones selected with the option *Show only*, then these categories will not be displayed. So the window in this particular example will only have two columns.

If you have a table with: Name, Title and Description and you set the Description to Hide, it means that Description will never be shown. But if the elements that you add to the window have other categories, these will still be displayed and the number of columns will increase accordingly (assuming that you don't use the show only option).

One more option of customizing the column view is the function **Limit View to Current Columns**, which you will find by right clicking the header of a column. When chosen, this mode will freeze the current columns view, meaning that no matter what element you add to the window, only the current categories (columns) will be shown.

It is also possible to set the width of the columns by dragging the right or left edge of the column. This applies to arbitrary open tables as well as saved column views.

Figure 5.15. Editing column view

It is also possible to hide a column from the current view. To do so, right click on the column that you do not want to be displayed and choose **This column (...)** > **Hide Column**.
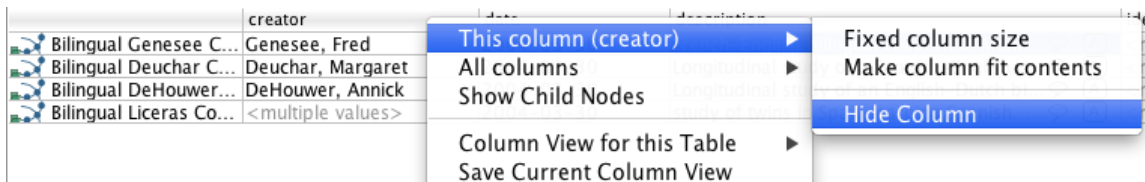
Figure 5.16. Hiding one column

> ⚠ **Note**
>
> The right click menu option **This column (...)** also serves to adjust the size of the columns. Through the secondary options (see figure above) you can either make the column size fix (**Fixed column size**) or to make the columns big enough in order for their content to fit (**Make column fit contents**).

Once you have hidden the unnecessary columns, you can save this view by right clicking any of the headers of the columns and choosing **Save Current Column View**. Give the column view a name and select **OK**.
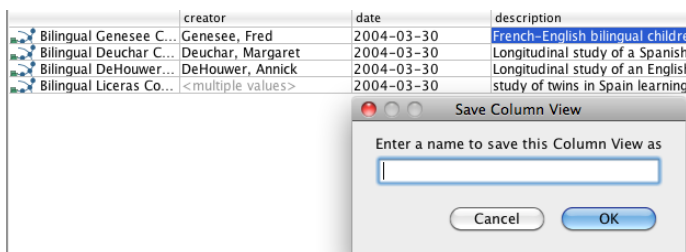
Figure 5.17. Saving Column View

Currently it is NOT possible to delete the column views that you create.

The view you have just saved is also applicable to other tables. Right click the header column of the new Table, select **Column View for this Table** and choose the view you want. By default, there are two views already defined and cannot be modified: **All** (i.e. all columns are displayed), and **Minimal** (i.e. only *Name*, *Title* and *Description* are displayed). It is important not to forget that each view you create initially is window specific. In this way you can use different views for different windows.

It is also possible to set your modified column view as the default column view that will appear in all tables you open. To do so, go to **Options** menu in Arbil toolbar, select **Column View For New Tables** and the preferred column view. This is now your default column view. (see number 1 in the figure below).
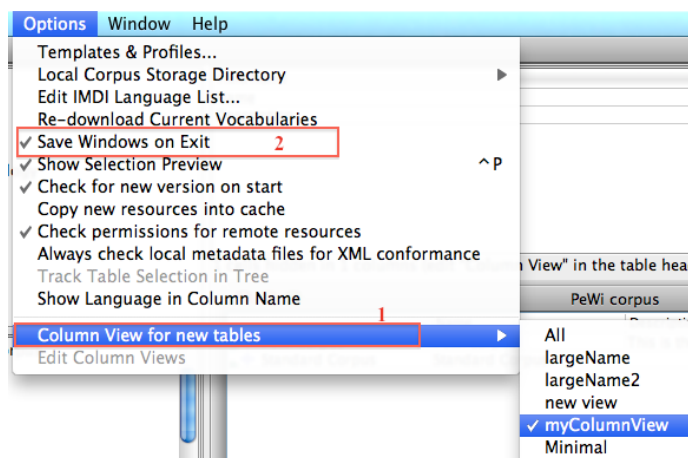
Figure 5.18. Setting new default column view

> **Note**
>
> Column views that have been modified and multiple open windows in the working area, either saved or unsaved, are available after closing Arbil once the programme starts again exactly in the same way they were before closing it provided the **Save Windows on Exit** is activated in the Options menu (see number 2 in the figure above).

5.8. Highlighting cells with the same value

If you want to see which cells in a table have the same value, right click the appropriate cell and choose **Highlight Matching Cells**. Cells with the same value will be highlighted. You can do this with more than one cell; each new group of matching cells will be given a different colour. To remove the highlight, right click on a cell and choose: **Clear Cell Highlight**.
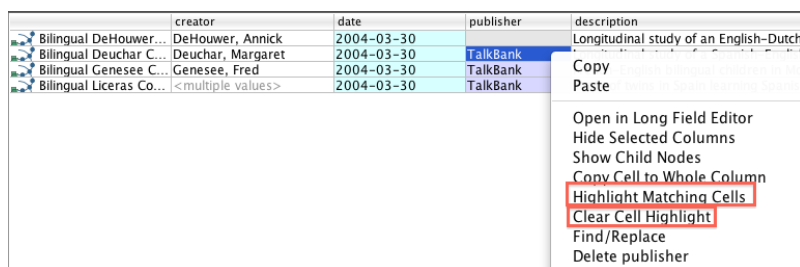
Figure 5.19. Highlighting cells

5.9. Track Table Selection in Tree

It may be the case that you have, in the main area of the application (i.e. the bottom right part in the figure below) an open table with many elements displayed. Given this high amount of resources/sessions, you may not be able to recognize which one you need, or else which is which. From the Options menu mark the option **Track Table Selection in Tree**: now if you select an item in the table, its corresponding location on the tree will be highlighted, thus enabling you to know which and where you required item is.
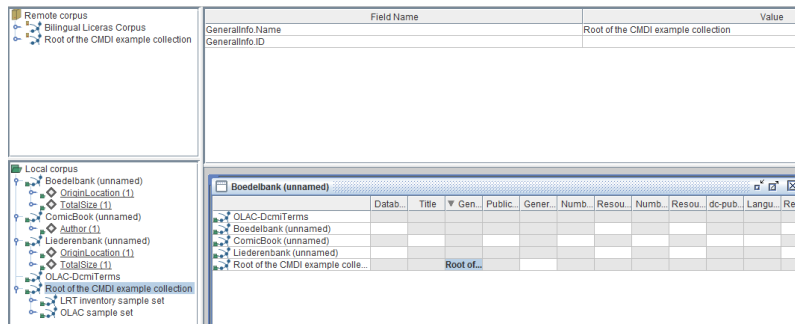
Figure 5.20. Track Table Selection inTree

Chapter 6. Searching

6.1. Searching Local Corpus

The search option allows you to find a value throughout either the Local Corpus or the Remote Corpus. To do so, right click on a given node you want to search and select **Search** (you might also choose a number of nodes by holding Ctrl and clicking on them). A window will appear in which you can specify where and what exactly you want to search e.g. within Corpus, Session, Catalogue, Actor etc.
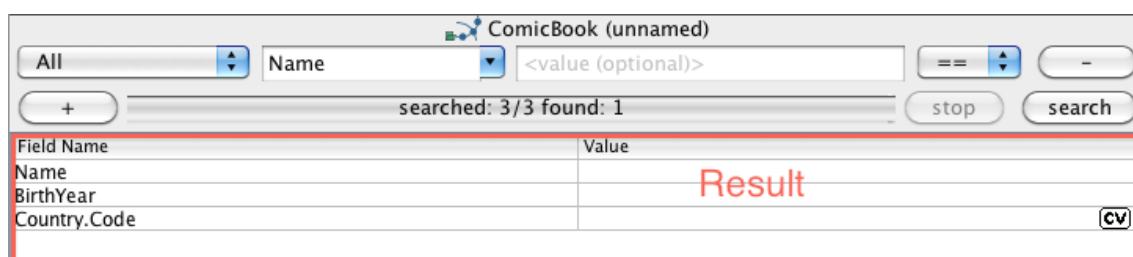
Figure 6.1. Selecting search locations

In the search window you can further specify in which column you want to search. By clicking on **<column (optional)>** a drop down list appears, containing pre-defined common values and previously entered column names; it is also possible to manually add column names to this list, or to remove arbitrary items by selecting the value and pressing the **delete** key on your keyboard; deleted values have to be retyped and searched for in order to appear in the drop down list again. After specifying the column (e.g. Name), give the value name you look for, e.g. "Blance", and click on **search**.

By selecting either *Equal* (== ▾) or *Unequal* (!= ▾), you will need to specify whether your search results should match the word you filled in or rather show you where this word does not appear.

6.2. Display of Search Results

Suppose the search we are going to make is: Creator (<column (optional)>) equals (==) Annick (<value (optional)>). After clicking on search, the results will be displayed in a table such as the one below. In case the content of one or more cell exactly matches the value you have searched for, the cell(s) will be highlighted.
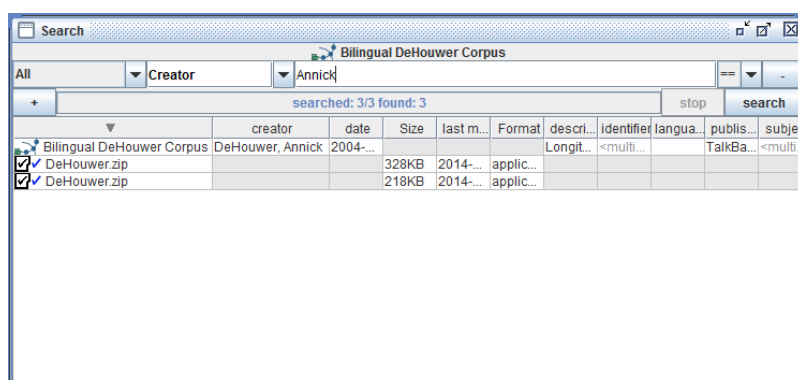
Figure 6.2. Display of Search Results

Chapter 7. Exporting and Importing data

7.1. Exporting data from Arbil

After creating new corpus branches or editing corpora from an already existing archive, you can save your files and upload them to the appropriate archive. In order to do so you need to export your files from Arbil. Select the modified nodes or the branch that contains the modified parts, right click on them and select **Export**. A dialogue box will pop up, in which you choose the location of the export and the name of the file. To complete the export, click on **Export Branch Destination Directory**.
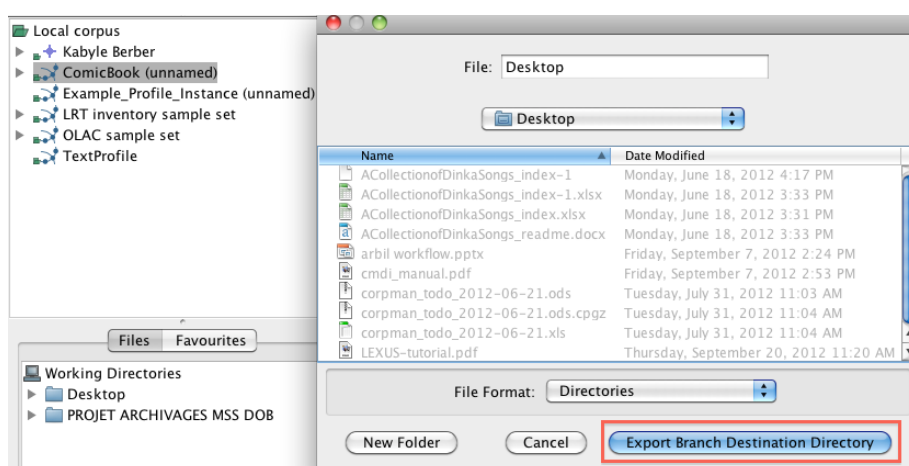
Figure 7.1. Exporting files

The following message may appear if the destination directory is not empty:

Figure 7.2. Export alert message

Selecting **OK** will lead you to a message window which is the same as the one seen above regarding the import into the Local Corpus (see Figure 7.3).
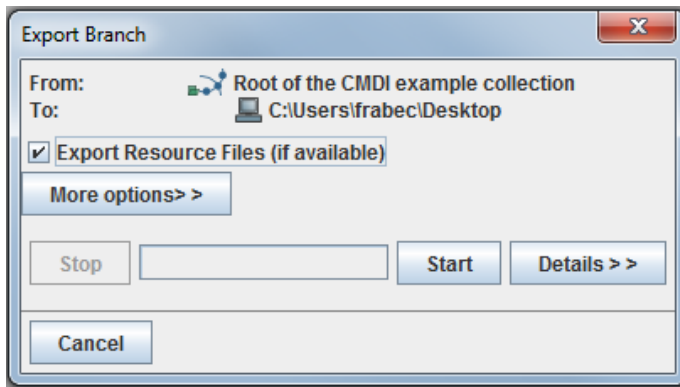
Figure 7.3. Export window with details

7.1.1. More export options

The export window offers advanced export options. These can be accessed by clicking the button **More options** (see Figure 7.4).
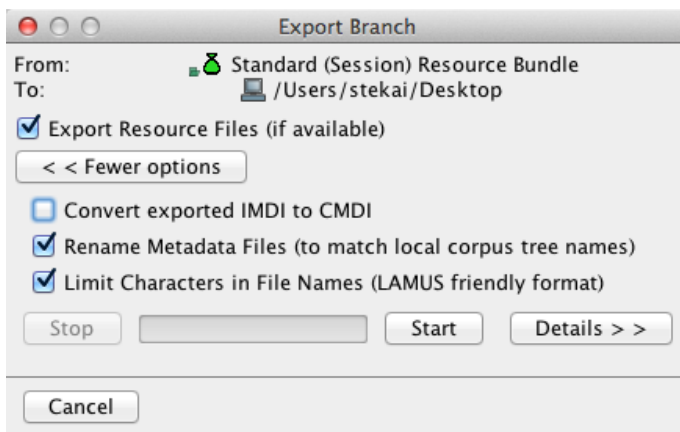
Figure 7.4. Export window - more options

The following options are available:

- **Convert exported IMDI to CMDI**: This option converts the exported IMDI files to the CMDI file type. This gives you the possibility to import the exported files to a CMDI archive.

- **Rename Metadata Files** : Rename files to match local corpus tree names.

- **Limit Characters in File Names**: This replaces all characters in the filenames which are not accepted by LAMUS. Currently accepted are digits, non-accented letters, dots (.), underscores (_) and hyphens (-). All other characters will be replaced.

7.1.2. Export process details

By selecting **Details** you can inspect the export process details including possible erros (see Figure 7.5).
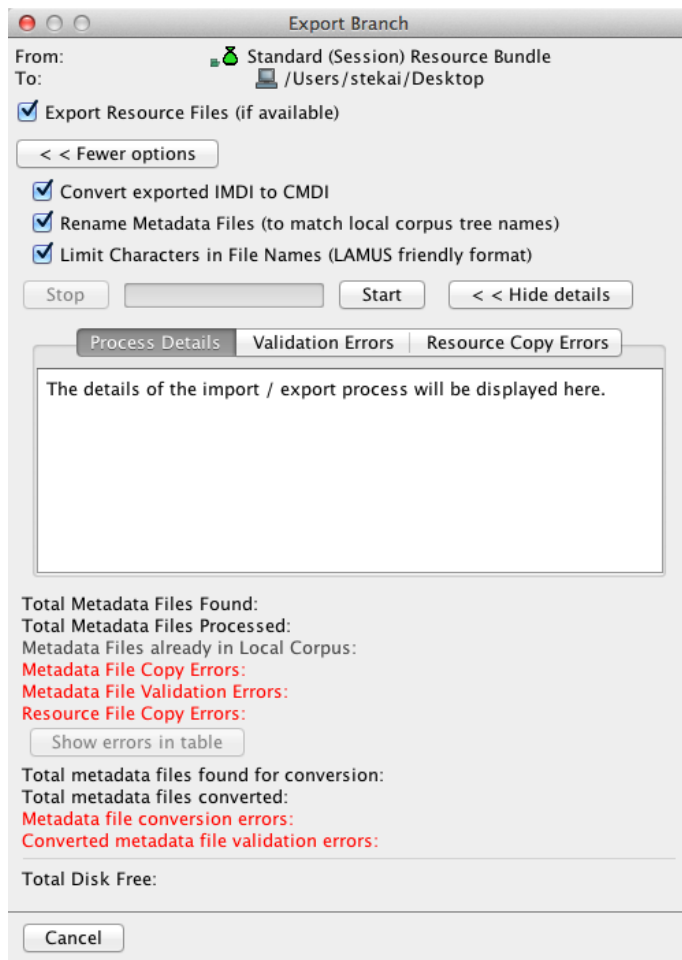
Figure 7.5. Export window - details

7.2. Importing data into Arbil

In Arbil there is a number of possibilities to import the data.

The first way to import data is from the Remote corpus to the Local corpus, either by right clicking on the required branch in the remote corpus and selecting **Import to Local Corpus**, or by dragging and dropping it.

If you want to import data from you local computer disk, or from the network, go to **File** menu, select **Import**, and then choose the required data from your local disk, or from the network; your data wil be imported at the root level.

In addition, at Local Corpus level, the context menu that comes out when right clicking one of the nodes includes further import options:

Import Branch (i.e. importing a piece of your corpus) gives you the ability to import part of your corpus from your disk to the Local corpus.

Reimport this branch (i.e. re-importing form the remote corpus): in order to make changes to a corpus you have already exported to the Remote corpus via Lamus, it is essential to reimport such corpus to the Local corpus and choose Yes to the message question "Overwrite?".

7.3. Using an External Editor

As is has already been noticed, it is possible to open, view and modify the data using external editors such as spread sheets or text editors. Overall the process is done by moving data from Arbil to the external editor, making the necessary changes, and then moving them back to Arbil.

However, possibilities of moving data back into Arbil are currently very *limited*. It is possible to do any of the following:

- Copy single field values (individual cells) from the external editor into Arbil;

- In case of the IMDI Session root node, save your data in *CSV format*, and then import them to Arbil by right clicking on the node where you want to copy them back and selecting **Import CSV**. This way you will be able to import the data in a table format; in any other case you will have to import the edited data column by column, but not all at once;

- If cells in a table corresponding to a single metadata file are copied, then the data will be in a format that can then be pasted back into any table in Arbil, provided that the data maintains the same formatting.

The following functions are currently *not available*:

- Pasting of multiple rows (i.e. data representing multiple nodes) into Arbil, even if the table in Arbil has an identical structure;

- Pasting of multiple values in general unless their structure corresponds to one of the options listed above;

- Importing CSV to any other type of node than the Session root.

To open the data in an external editor (i.e. Ms Excel), select and copy the data from the table, then paste them in the editor you wish to use.
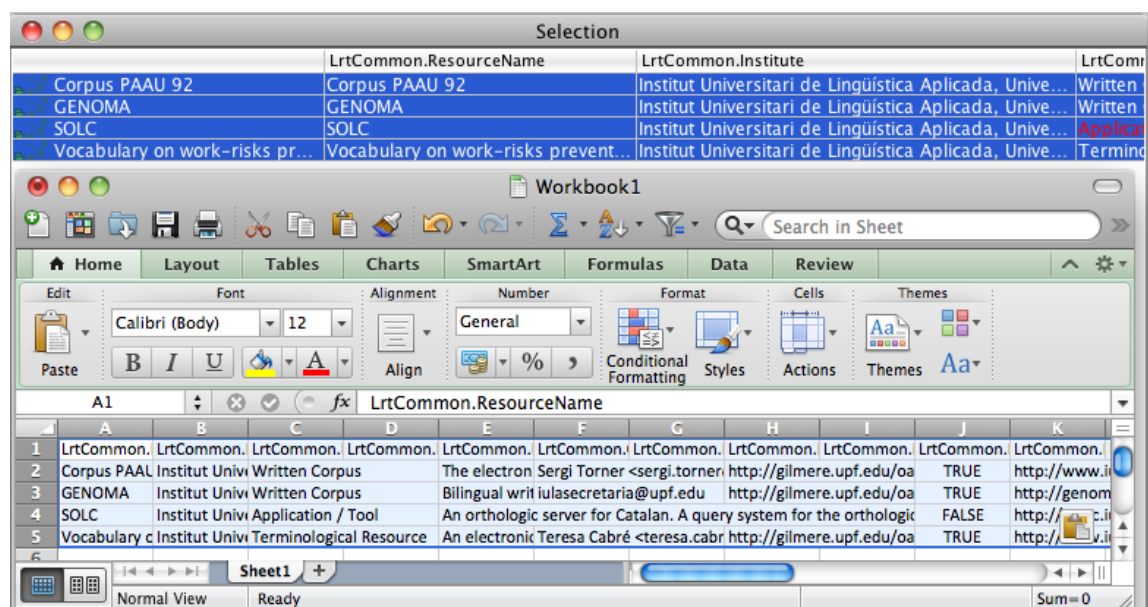
Figure 7.6. Editing data in an External editor

7.4. External application

By selecting *Open in External Application* you can view and edit resources and metadata in the application that your computer chooses by default, such as a media player for a video file, or a text editor for a text file.
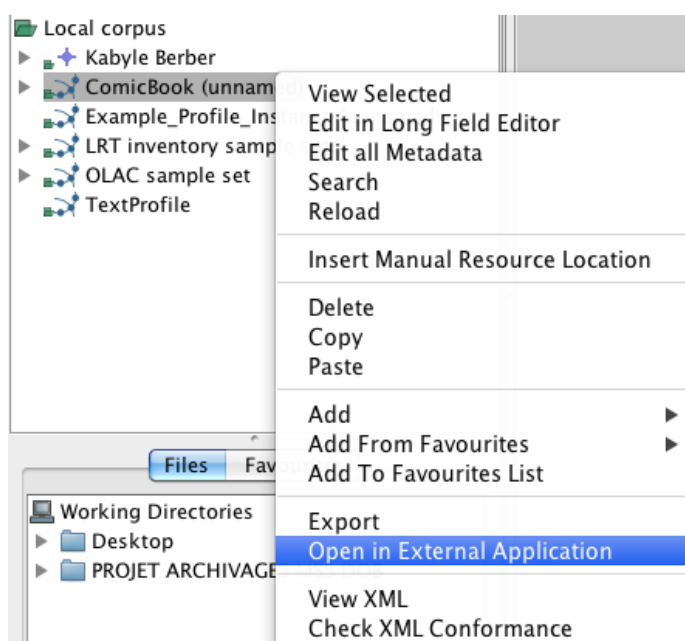
Figure 7.7. Viewing data in external application

Chapter 8. Exiting Arbil

Before exiting Arbil make sure that you have saved all your changes. This can easily be done by the key short-cut CTRL/CMD+S.

You can exit Arbil by closing the main window, or through *File -> Exit* on Windows or Linux, or Arbil *-> Quit Arbil* on Mac OS. In case of unsaved changes a message will appear and ask if you wish to *Save Changes Before Exiting*; choose *Yes* or *No*, Arbil is now closed.

Chapter 9. Shortcut Keys

In Arbil you can use the following shortcuts:

Table 9.1. General shortcut keys

CTRL/CMD+A	Select all
CTRL/CMD+F	Find & Replace
CTRL/CMD+S	Save
CTRL/CMD+W	Close topmost window
CTRL/CMD+SHIFT+W	Close all windows
CTRL+TAB	Next window
CTRL/CMD+SHIFT+TAB	Previous window
CTRL+P/CMD+P	Hide/Show the selection preview
CTRL++/CMD++	Increase the font size used in the application
CTRL+-/CMD+-	Decrease the font size used in the application
CTRL+0/CMD+0	Reset the font size used in the application to the default value
ENTER (in the tree)	Open the current tree selection in a table
DELETE (in the tree)	Delete the selected node from the tree

Table 9.2. Table short cut keys

CTRL/CMD+A	Select all
CTRL/CMD+C	Copy
CTRL/CMD+V	Paste
CTRL/CMD+Z	Undo
DELETE	Delete the contents of the current table cell and start editing
ESC	Cancel editing of the current cell
CTRL+ENTER	Edit a cell in a new window

Table 9.3. Tree short cut keys

DELETE	Delete the selected node in the tree

Table 9.4. Long field editor short cut keys

CTRL+UP	Go to previous field
CTRL+DOWN	Go the next field
CTRL+LEFT	Go to previous tab
CTRL+RIGHT	Go to next tab

Chapter 10. Technical details

10.1. Custom File types Configuration

It is possible to override the default set of file types that the type checker should recognize as archivable resources.

To do so, place your custom filetypes.txt configuration file in the root of your local storage directory (to find out what that is, select Local Corpus Storage Directory from the Options menu). It will then automatically be installed in the type checker as soon as type checking takes place for the first time during a session. You will need to restart Arbil if it is already running.

If you have already typechecked files within your working directories when you replace the configuration, Arbil will still remember the old status for these files. To make Arbil forget about this, you can delete the following three files from the root of your local storage directory:

- knownMimeTypesV2

- md5SumToDuplicates

- pathToMd5Sums

Arbil will now recheck every file within its working directories, so depending on how much is in there, this might take some time.

For information about obtaining or creating a custom file types configuration, please contact your administrator.

10.2. Templates

Templates are currently in development, but they will be used to customise the components added via the "Add" menu.

To create a new template, create a directory in the ".arbil/templates" directory. This is where all the files relevant to that template will be stored. This directory will now be selectable via the "options/templates" menu.

Each template can have its own XSL which is used via the "view formatted" menu. The XSL must be in the selected templates directory and be called "format.xsl". For example: .arbil/templates/{template name}/format.xsl. If there are any additional files required by the XSL they can be placed in the same directory as the XSL file.

10.3. Plugins

Plugins add extra functionality to Arbil. They are provided as '.jar' files that can be placed in a special location where Arbil will detect them and add them as items to the 'Plugins' menu (which is hidden unless plugins have actually been detected).

To install a plugin, place the .jar file in one of the following directories:

- ~/ArchivingPlugins (Linux or MacOS)

- %APPDATA%/ArchivingPlugins (Windows)

- %USERPROFILE%/ArchivingPlugins (Windows)

- <ARBIL STORAGE DIRECTORY>/ArchivingPlugins

If none of the locations already exist, you can create one. Some options will not be available on all operating systems. ~ denotes the user's home directory on Unix systems. %APPDATA% and %USERPROFILE% are system environments available in Windows. Use the 'run command' option (Windows key + R) to open locations that contain the environment variable.

<ARBIL STORAGE DIRECTORY> is the directory called .arbil that contains the configuration files and by default is the parent directory of the working files directory.

After installing the plugin you should restart Arbil for it to become available in the Plugins menu of the menu bar.

10.4. Custom logging configuration

If present, a logging configuration will be read and applied from a file called 'logging.properties' in the application configuration directory ('.arbil'). An example of such a configuration is given below.

The logging output (visible in the log console of the Help menu) can be observed from different levels of detail. Such levels are: **ALL**, **FINE**, **INFO**, **WARNING**, **SEVERE**. Note that the log console is visible only when the application is installed, i.e. NOT with the webstart version.

```
handlers= java.util.logging.FileHandler, java.util.logging.ConsoleHandler

# Default global logging level.
.level = ALL

#########################################################
# Handler specific properties.
# Describes specific configuration info for Handlers.
#########################################################

# default file output is in user's home directory.
java.util.logging.FileHandler.pattern = %h/arbil.log
java.util.logging.FileHandler.level = WARNING
java.util.logging.FileHandler.formatter = java.util.logging.SimpleFormatter

# Limit the message that are printed on the console to INFO and above.
java.util.logging.ConsoleHandler.level = FINE
java.util.logging.ConsoleHandler.formatter = java.util.logging.SimpleFormat


#########################################################
# Facility specific properties.
#########################################################

java.level = WARNING
javax.level = WARNING
sun.level = WARNING

nl.mpi.arbil.level = FINEST
nl.mpi.flap.level = FINEST
nl.mpi.level = FINEST
```

Appendix A. CMDI ecosystem

CMDI - "Component Metadata infrastructure" – is a metadata framework in which metadata blueprints can be described and reused.

Metadata is data about data, giving various kinds of information about the data contained, e.g. about protagonists, places, the date of gathering data etc. To manage this metadata in a diligent way certain ready-made options for description can be used which facilitate the production of well-ordered information.

In addition, to create metadata (components) from scratch, CMDI allows you to use existing description options –blueprints in the form of 'profiles and components' (see Figure A.1 below). These existing blueprints can be edited in the component registry according to individual needs, if necessary.

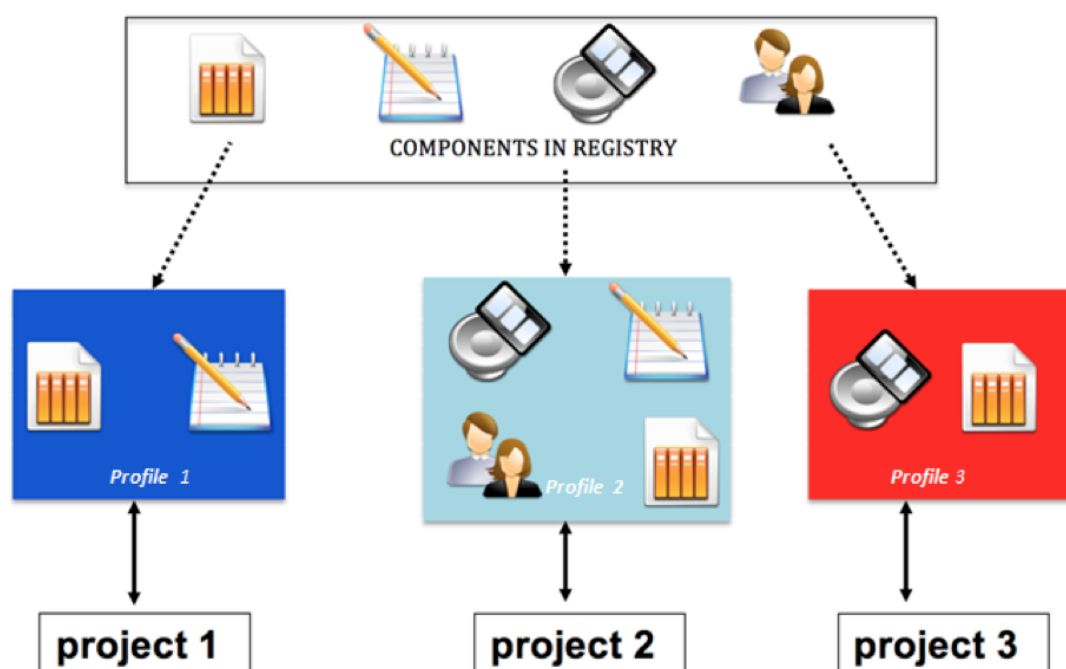Furthermore, Arbil also allows editing the content of metadata descriptions.

Figure A.1. Building profiles from components according to the projects' individual demands with CMDI in the component browser.

CLARIN has established this infrastructure of ready-made building blocks (components) to overcome "dispersion of metadata in diverse formats" (see CLARIN website [http://www.clarin.eu/node/3219] for more information about CMDI): through the use of standardized building blocks metadata will not be subject to arbitrary and very diverse, individually preferred ways of describing data about data (=metadata). All the necessary information can be extracted easily from the standardized metadata-descriptions without having to study the data themselves and first understanding the individual way authors of metadata have set up their descriptions. Through standardization, working with metadata is thus made easy for authors and users. All in all, the potential of sharing and storing self-designed components and profiles for reuse by others facilitates flexible, but structured metadata building and search among various language resources.

Another advantage of CMDI is the possibility of linking elements in components and profiles to the ISOcat concept registry [https://catalog.clarin.eu/isocat/interface/index.html]. This registry is a central framework for gathering concepts in various domains and has special profiles for metadata related categories. By creating links to this registry users searching for certain kinds of data in ISOcat will be enabled to find all data belonging to different authors, thanks to standardized metadata descriptions and to the central registry. It is thus highly recommended to browse through ISOcat and to create links to it when working with CMDI, which can be done from inside the Component Registry.

The last significant CLARIN-related instrument to be mentioned here is the VLO [http://catalog.clarin.eu/vlo/] (Virtual Language Observatory). This website gives an overview of metadata from a certain amount of registries covering a large number of institutions which make use of CDMI. Additionally, all metadata which has been created with the help of CMDI is available in the several records provided by the institutions. This service facilitates the search for, and the exchange of, data with the help of metadata by researchers' communities in the field.

A.1. Creating profiles and components in the Component Registry

For building, searching and using profiles and components a tool has been developed, which is the Component Registry. The component registry [http://catalog.clarin.eu/ds/ComponentRegistry/] (and its manual [http://www.clarin.eu/node/3223]) can be found on the CLARIN website.

This short introduction of CMDI will outline the most basic features for using the infrastructure.

A.1.1. Login

To login to the Component Registry you have to choose one's organization from the list or opt for "Clarin Guest Accounts". To log out simply end the browser session.

A.1.2. Features of the Component Registry

The first way of using CMDI is browsing through registered components and profiles. The second way consists of editing and creating components or profiles.

Registering and storing components or CMDI profiles are further functions of the component registry.

To illustrate the workflow in the component registry please consult the following Figure A.2:

1) Examine existing profiles/components (browser)

2) Create own components if necessary/profiles and combine existing ones (editor)

3) Link to datacategories in ISOcat (editor)
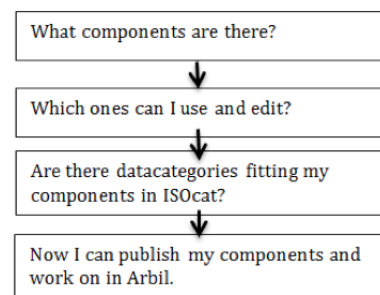
4) Publish and instantiate Arbil (editor and Arbil)

What components are there?
↓
Which ones can I use and edit?
↓
Are there datacategories fitting my components in ISOcat?
↓
Now I can publish my components and work on in Arbil.

Figure A.2. Workflow from CMDI to Arbil

A.1.2.1. Component Browser

The start screen of the Component Registry [http://catalog.clarin.eu/ds/ComponentRegistry/#] is called the component browser, which is the entry point for using other functions (editing and importing) (see Figure A.3). First, though, the user can browse for either profiles or components. Search-terms can be typed in a field at the top on the right to filter out non-matching items. Additionally the user can switch between public and (unpublished, private) user workspaces. For more information on individual profiles or components, one can select the entity in question and details will appear in the panel below. Such panel can be optionally displayed in an interactive default view (to explore structure more in depth) or in an xml-view of the actual component assessed.

Right-clicking on items in the list leads to options such as downloading profiles or components as xml representations or downloading the respective xsd. The latter can be loaded into Arbil for usage before publication (see section ' Section 3.5'), and suits further use by other tools within CLARIN.

On the top-left of the Component Registry you will find three options: the first option (**Create new**) allows creating a new profile or component in a private workspace. If you choose the second (**Edit as new**), a selected item can be opened in the editor after authentication of the user, depending on whether it is a copy of a public space or an item in the users' workspace. The last option (**Import**) leads to an import form, which will not be discussed in this chapter.
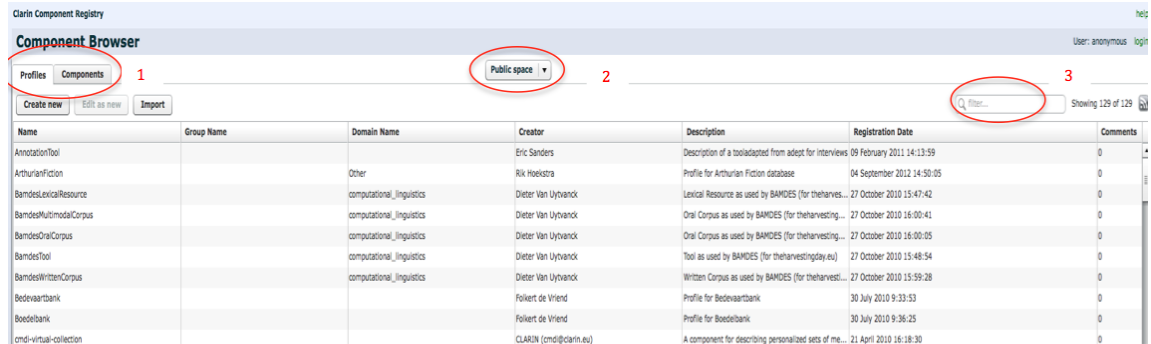

Figure A.3. Component Registry window and functions offered: The first red demarcation shows the option to choose between profiles and components (1). The second demarcation shows the option to choose for public or private space (2). The third demarcation highlights the filter option field (3).

A.1.2.2. Component Editor

With the help of the component editor the user can modify a profile and/or its components (see Figure A.4). You can access it - after having selected the profile to modify - via the option **Edit as new** mentioned above. A page like the one shown in the figure below will be shown.

Figure A.4. Component editor. The first red demarcation indicates where you find the main menu bar in the component editor (1). The second demarcation shows where information is found about whether a component or a profile is edited (2). The third one shows the option to add attributes (3). In the fourth demarcation the field for editing elements is stressed (4). The last highlights the option to choose from public or private space to drag existing components into the editor (5).

To reuse existing components (Edit (as new) –function) select an existing component from the list below, drag it over the editing panel. Here options will be highlighted where dropping the component is possible. If the component already exists, at this point a reference will be automatically generated to it. Be aware that components which you you have not published yet cannot be used by anyone else. Instead referenced components which you create and publish can be reused by others (when viewing profiles these components are marked by a blue frame in the list of the lower part of the screen). The green buttons indicate options for adding new components ("+component", "+element", "+attribute"), so that newly created fields can be worked on. Up and down arrows allow to reorder elements and components.

The "save"-function can be used to submit and overwrite existing profiles or components in private workspaces. Existing profiles or components can be used as a basis for new profiles or components without overwriting the existing. In this case the user can apply "save as new" in their private workspaces and work on the profiles and components to adapt them to their needs. The function "publish in public space" is to submit a profile or component as a building-block accessible for others, finalizing the modeling phase of that particular profile or component since after publication, a profile or component cannot be changed anymore. To return to the component browser and discard changes, press "cancel".

For more detailed information about more elaborate editing processes (type, import, concept links, action links, comments and priority display) have a look at the CMDI metadata [http://tla.mpi.nl/cmdi-metadata/] page on the TLA website, the CLARIN website [http://www.clarin.eu/node/3219] - where you can access the Component Registry Documentation [http://www.clarin.eu/content/component-registry-documentation] - or the extensive CMDI manual [http://tla.mpi.nl/cmdi-metadata/] of the TLA.